OSU WPL 31 (1985) 1-63

English Adverb Placement in Generalized Phrase Structure Grammar*

Belinda Lea Brodie

1. Introduction

1.1 Objectives

While there have been many works in the last couple of decades dealing with the semantics of English adverbs, few syntactic analyses have been presented. The syntactic analyses which have been proposed have generally been motivated by semantic, rather than syntactic, considerations. The methodological approach taken in these works is to assume or argue for a particular semantic treatment of adverbs, and then provide a syntactic analysis compatible with this semantic treatment. This approach has led, I believe, to incorrect conclusions concerning the placement of adverbs. While I agree with the assumption underlying this approach (i.e that there is a correspondence between syntactic structure and semantic interpretation), I object to the priority given semantic considerations and the disregard for syntactic evidence.

The purpose of this work is to provide an analysis of adverb placement which gives priority to syntactic evidence, and which, moreover, accounts for a much wider range of data than previous analyses. In this work, it is assumed that each syntactic rule is associated with a particular semantic rule. However, it is the syntactic rule which limits the set of possible semantic rules, not vice-versa. I believe that this approach is to be preferred in the case of adverb placement, because there is, at least for some types of adverbs (i.e. evaluative, modal, temporal, and frequency adverbs), evidence which supports a unique syntactic treatment, but no evidence which requires a unique semantic analysis. In these cases it is the syntactic rule which will limit the possible semantic analyses. In cases where there is no evidence to decide between two or more alternate syntactic analyses (i.e. VP adverbs), semantic considerations should, of course, be used to determine the correct syntactic analysis, if possible.

In succeeding sections I will propose analyses of the placement of evaluative, modal, temporal, frequency, and VP adverbs. These analyses are given within the framework of Generalized Phrase Structure Grammar (GPSG). GPSG is a monostratal theory of syntax which is preferable to most other current theories of syntax on two grounds:

1. It is more restrictive, in terms of generative capacity. It has the generative capacity of a context-free phrase-structure grammar, rather than that of a more powerful type of grammar.
2. It is associated with a formal semantics (Montague Grammar) the properties of which are well-defined, and which is restricted by the requirement that the semantics be rule-to-rule.

It is shown that within GPSG it is possible to give a small number of syntactic rules which account for all the data accounted for by previous analyses of adverb positions, as well as a good deal of data not previously considered or accounted for. It is possible to show that these syntactic rules will allow for rule-to-rule translations which will yield the proper semantic results. In this paper, I have been able only to provide the semantic rules and give some indication of how certain aspects of the semantics of adverbs are accounted for by this treatment.

Not only is the analysis to be presented preferable to previous analyses because of its simplicity and scope, but also because this analysis, in conjunction with independently motivated aspects of GPSG, accounts for observations concerning adverb stranding which have not previously been given an adequate treatment. The only previous treatment which is close to being observationally adequate is one in which the notion of "trace" is necessarily referred to in a surface filter constraint (cf. Sag (1978, 1980)). The analysis presented here accounts for adverb stranding data without making reference to traces and is, therefore, consistent with Jacobson's (1982: 207) tentative claim that "no constraint in the grammar can explicitly mention gaps." It is significant that this claim can be maintained with respect to adverb stranding within a framework which is already more restrictive than most other current syntactic theories.

1.2 A brief introduction to GPSG

A GPSG consists of two parts--the actual grammar, which includes the set of phrase-structure rules of the language, and the metagrammar, which consists of rules and principles that characterize the phrase-structure rules and express generalizations between rules.

A phrase-structure rule consists of three parts: a rule number, a syntactic rule, and the semantic rule associated with the syntactic rule. In the PS rule below, for example, the rule number is 2, the syntactic rule is VP -> V VP and the semantic rule is V'(VP').

3. <2, [VP -> V VP], V'(VP')>

Rule numbers are used as subcategorization features on the lexical category node introduced by the rule. Thus, the PS rule above is an abbreviation for the rule in 4.

4. <2, [VP -> V VP], V'(VP')>
[2]

The phrase-structure rules are characterized by two types of rules of the metagrammar and by feature instantiation principles. Immediate Dominance (ID) rules express possible immediate dominance relations. The ID rule below, for example, states that A may immediately dominate B, C, and D.

5. A -> B,C,D

The immediate dominance relations expressed by any phrase-structure rule must be identical to immediate dominance relations expressed by

one of the ID rules. Linear Precedence (LP) rules express the ordering relationships which must hold between sister nodes. The rule below, for example, states that B must precede C in any phrase-structure rule in which B and C are sisters.

6. $B < C$

Each phrase-structure rule must be consistent with every LP rule of the metagrammar.

Feature instantiation principles govern the distribution of features. The Head Feature Convention, for example, ensures that the head of a phrase has head features identical to those of its mother. Every phrase-structure rule must be consistent with the Head Feature Convention, and all other principles of feature instantiation.

The metagrammar also includes a type of rule, known as a metarule, which does not characterize phrase-structure rules, but instead expresses implicational relationships between ID rules. The metarule below, for example, states that for every ID rule in the grammar in which A dominates some finite set of category symbols X, there is another rule in which A dominates this same set of symbols and also dominates B.

7. $\langle n, [A \rightarrow X], (F')\rangle \Longrightarrow \langle [A \rightarrow X, B], B'(F')\rangle$

This rule also states that the semantic interpretation of the rule on the right will be the result of applying the semantic value of B to the semantic value of F, which is a variable ranging over the interpretations of rules characterized by the syntactic rule on the left of the arrow. The rules related by a metarule will have the same rule numbers.

In earlier versions of GPSG, metarules expressed implicational relationships between PS rules, rather than ID rules. In quoting some earlier works, I will give the rule in terms of PS rules rather than ID rules, but, in every case, the metarule could have just as well been given in terms of ID rules.

I will use the slashing metarule, which Gazdar (1982) defines as follows:

> Let G be the set of <u>basic</u> rules (i.e. the set of rules that a grammar not handling unbounded dependencies would require). For any syntactic category B, there will be some subset of the set of the nonterminal symbols V_N each of which can dominate B according to the rules in G. Let us call this set V_B ($V_B \subset V_N$). Now, for any category B ($B \in V_N$) we can define a (finite) set of derived rules D (B,G) as follows:
>
> $$D(B,G) = [a/B \rightarrow o_1 \ldots o_i/B \ldots o_n] : [a \rightarrow o_1 \ldots o_i \ldots o_n] \in G \;\&\; 1 < i < n \;\&\; a, o_i \in V_B$$

The slashing metarule is described in this passage as applying to PS rules (basic rules) to allow other PS rules (derived rules). I will sometimes refer to the slashing metarule applying to basic phrase-

structure rules to yield derived phrase-structure rules, but, in every case, the slashing metarule could just as easily have applied to ID rules to allow new ID rules. To give an example of the application of the slashing metarule: Given the basic rules in 8,

8. VP -> V VP
   VP -> V S
   S -> NP VP

the derived rules in 9 will be allowed by the slashing metarule.

9. VP/VP -> V VP/VP
   VP/VP -> V S/VP
   S/VP -> NP VP/VP

In Gazdar and Pullum (1982) the work done by Gazdar's slashing mechanism is carried out by a feature slash. However, whether the slashing mechanism or the feature slash is used is irrelevant to the analyses I will propose.

It is important to point out that, although in giving rules I have used ADVERB (eg. S -> ADV S) rather than ADVERB PHRASE, replacing ADV in these rules with ADVP would require only a slight revision in the statement of rules. In the rules using ADV, I have made use of the lexical status of ADV to subcategorize adverbs with respect to their sisters so that differences in positions of occurrence could be accounted for. If ADV is replaced by ADVP in the proposed rules, this approach is no longer possible. Instead, we must distinguish various categories of ADVP which dominate different lexical categories of adverbs and give rules for the placement of ADVP, allowing different ADVP categories to occur in different rules.

Finally, it should be noted that, for convenience sake, I have replaced all references to V', N'' etc. with VP, NP etc. In the version of GPSG which I adopt both matrix and embedded VP's are assigned one bar, thus the use of VP is not problematic. My use of S corresponds to V'', the maximal projection of V. In the semantic rules which I give I use the type assignments of Klein and Sag (1982) and follow their convention of not mentioning intentions in the semantic translation; however, when quoting rules, I give the semantic translation as it originally appeared.

FOOTNOTE

*I would like to thank Arnold Zwicky, Mike Geis, and especially David Dowty for their helpful comments and criticism on this work. I am of course solely responsible for all errors.

2. Previous Analyses of Adverb Positions in English

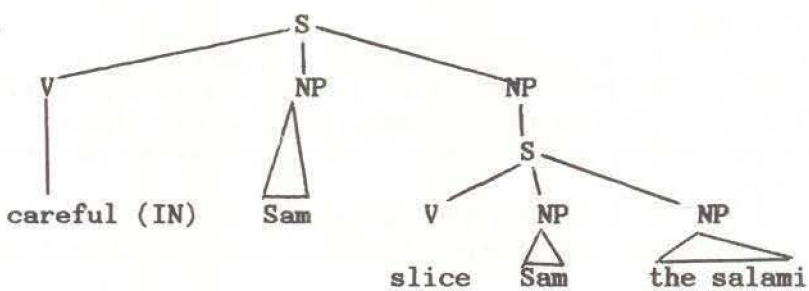
2.1 Lakoff (1965, 1970)

In work by generative semanticists, it was assumed that adverb placement was accounted for by one or more transformational rules. Generative semanticists used the similarities in the selectional restrictions of adverbs and adjectives to argue that adverbs should be transformationally derived from adjectives. Lakoff (1965, 1970) argued that manner adverbs should be derived transformationally from their corresponding adjectives. Similarly, Schreiber (1971) claimed that sentence adverbs should be derived by transformation from their corresponding adjectives.

Lakoff claimed that sentences containing manner adverbs should be derived from the same underlying structures as sentences containing the corresponding adjectives. Both 1 and 2 below were derived from the underlying structure in 3 by means of an adverb lowering transformation. This transformation deletes the occurrence of Sam in the highest clause and moves careful into the lower clause, adding ly onto it.

1. Sam sliced the salami carefully.
2. Sam was careful in slicing the salami.

3.



Lakoff argued for this analysis on two grounds. First of all, he claimed that the elimination of the category Manner Adverb from the set of underlying categories of English would result in simplification of the grammar. The base component is simplified, but the transformational component is complicated by the addition of a rule. Whether or not Lakoff's analysis simplifies the grammar as a whole cannot be determined unless the values of the various elements of the grammar are specified.

Secondly, he claimed that his analysis would eliminate redundancy in the statement of selectional restrictions. In particular, if underlying structures such as 3 are adopted, the anomaly of sentences such as 4 and 5 will be accounted for by the selectional restrictions between underlying subjects and adjectives, and a second set of selectional restrictions between underlying subjects and adverbs need not be included in the grammar.

4. Moss hangs from trees recklessly.
5. Moss is reckless in hanging from trees.

It is, of course, possible to avoid such redundancy without resorting to a transformational derivation of adverbs. In the analysis to be presented here, it will be assumed that meaning postulates (cf.

Dowty (1980)) account for the logical entailment between sentences such as 4 and 5. Once such meaning postulates are adopted, the semantic incoherence of sentences such as 4 will follow from the semantic incoherence of sentences such as 5.

The generative semantics' notion that adverbs are transformationally derived was rejected by transformationalists such as Bowers (1969) and Jackendoff (1972, 1977). They noted that the same arguments which Chomsky (1970) gave in favor of a lexical, rather than transformational, treatment of "derived" nominals also apply in the case of adverbs.

It is now generally accepted that derivational rules, such as the rule creating adverbs from adjectives, apply only in the lexicon. In the analysis to be presented, it will be assumed that a lexical rule (in the sense of Dowty (1978)) derives adverbs from adjectives. It is characteristic of lexical rules that unprincipled exceptions to the rule occur. Schreiber (1971) notes "accidental gaps" such as nicely and improbably, which he considers to be possible but nonoccurring sentential adverbs.

2.2 Jackendoff (1972, 1977)

One of the few syntactic analyses of adverbial positions is that presented in Jackendoff (1972) and slightly revised in Jackendoff (1977). Jackendoff deals with two main classes of adverbs, those traditionally called sentential adverbs and those known as VP adverbs or predicate modifiers.

Jackendoff's analysis is intended to account for the following claimed generalizations concerning the positions of sentential and VP adverbs.

6. A sentential adverb may occupy any position in which it is a daughter of the node S.
7. A VP adverb may occur in any position in which it is a daughter of the node V''.

It is important to note that Jackendoff's notion of V'' is distinct from the notion of V'' used in the version of GPSG adopted here. In Jackendoff's analysis V'' dominates V' and optionally dominates constituents for which the verb in V' is not strictly subcategorized. V' dominates the verb and any constituent for which the verb is strictly subcategorized. Thus, Jackendoff's claimed generalization in 7 predicts that VP adverbs will not intervene between a verb and any constituent for which the verb is strictly subcategorized, since VP adverbs are always daughters of V'', and not V'. Counterexamples to this prediction will be discussed later.

Jackendoff's analysis makes use of the "transportability convention" of Keyser (1968) in order to capture the claimed generalizations in 6 and 7. The transportability convention permits a constituent marked as transportable "to occupy any position in a derived tree so long as the sister relationships with all other nodes in the tree are maintained, that is, as long as it is dominated by the same node." (Jackendoff 1972, p. 67). Jackendoff (1977) claims that sentential adverbs and VP adverbs are transportable constituents. Sentential adverbs will be generated as daughters of S by phrase-structure rules and the transportability convention will allow the sentential adverb to move to any position as long as it remains a daughter of S. VP adverbs

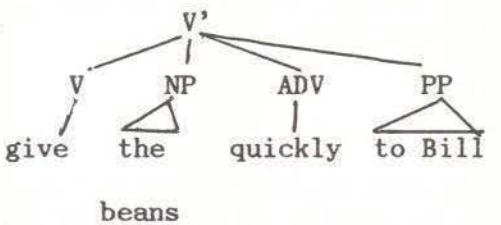
will be generated as daughters of V'' by phrase-structure rules and the transportability convention will allow the VP adverb to move to any position as long as it remains a daughter of V''.[1]

It must be pointed out that Jackendoff was not necessarily assuming that these generalizations about adverb positions will hold at the surface structure level. This becomes clear when Jackendoff discusses a class of examples which are problematic for his generalization concerning VP adverbs. The problematic examples are sentences such as 8 in which the VP adverb precedes a PP for which the verb is strictly subcategorized.

8. John gave the beans quickly to Bill.

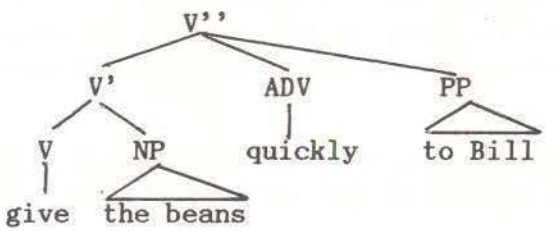
Since give is strictly subcategorized for the PP, the PP will be generated by Jackendoff's phrase-structure rules as a daughter of V'. But Jackendoff's generalization in 7 predicts that the adverb will be a daughter of V'', not V'. Jackendoff considers two solutions. The first solution which he considers is to generate the adverb as a daughter of V'' and then lower it by a transformational rule into position as a daughter of V', yielding the surface structure in 9.

9.



If this solution is adopted, the generalization in 7 is met at the deep structure level, but not at the surface structure level. The other solution which Jackendoff considers is to generate the PP as a daughter of V' but then to raise the PP into position as a daughter of V'', giving the surface structure in 10. On this account, the adverb is a daughter of V'' at both the deep and surface structure levels. Thus, if this solution is adopted, the generalization in 7 is met at both the deep structure and surface structure levels.

10.



Jackendoff does not decide between the two solutions; thus it is not clear whether or not he intends his generalizations to be generalizations about surface structure. If Jackendoff's generalizations concerning positions of sentence and VP adverbs are true generalizations about surface structure positions, then they are easily translated into a monostratal theory in which immediate dominance and linear precedence relations are stated separately (cf. Gazdar and

Pullum (1981)). In immediate dominance (ID) rules, the daughter constituents are unordered with respect to one another. In the ID rule in 11, for example, B, C, and D are unordered. Linear precedence (LP) rules express linear ordering relations between sister constituents. Rule 12, for example, states that B will precede C when they are sisters. The set of phrase structure rules of the grammar includes all rules consistent with some ID rule and every LP rule.

11. A --> B, C, D
12. B < C

In a grammar in which immediate dominance and linear precedence relations are expressed by distinct rules, Keyser's notion of a transportable constituent corresponds to a category which does not appear in any linear precedence rules. Such a category will be unordered with respect to other categories and may, therefore, either precede or follow any of its sister constituents. If Jackendoff's generalizations are meant to hold at the surface structure level, they can be expressed in a version of GPSG which adopts Jackendoff's assumptions about constituent structure by allowing metarules 13 and 14, and by <u>not</u> including the category ADVERB in any LP rules. (I have omitted the semantic translations in 13 and 14, since it is the syntactic generalizations that are at issue here.)

13. <1, V''-> X> ==> <V''-> X ADV>
14. <2, S -> X> ==> <S -> X ADV>

Metarule 13 states that for any ID rule which expands V'' as a finite set of categories X, there will be another ID rule which expands V'' as X plus the category ADVERB. Metarule 14 states that for any ID rule which expands S as X, there will be a rule expanding S as X plus ADVERB. Since ADVERB will not be ordered with respect to any of the categories in X, the metarule in 13 will allow adverbs in the lexical class 1 (i.e. VP adverbs) to appear in any position as daughter of V''. The metarule in 14 will allow adverbs in the lexical class 2 (i.e. S adverbs) to appear in any position as daughter of S.

These two metarules will account for Jackendoff's generalizations in 6 and 7 assuming they refer to surface structure. Unfortunately, this simple analysis cannot be maintained. Jackendoff's generalizations, when considered to apply at the surface structure level, lead to incorrect predictions and rely on unmotivated assumptions about constituent structure.

In order to account for sentences such as 15, for example, Jackendoff must assume that the first auxiliary, but not subsequent ones, is a daughter of S.

15. John will probably leave in the morning.

The only motivation he gives for assuming that the first auxiliary is a daughter of S is that adopting this structure allows the positions of S adverbs in sentences such as 15 to be accounted for by his analysis of adverbs:

> Jackendoff (1972) gives evidence that the first auxiliary is a daughter of S, but that subsequent auxiliaries are not daughters of S. The evidence is that sentence adverbs such as frankly, probably, and evidently occur in all possible positions as daughters of S - initial, final with comma intonation, and before the auxiliary. They also occur after the first auxiliary, but not after subsequent ones.
> (Jackendoff (1977:48)

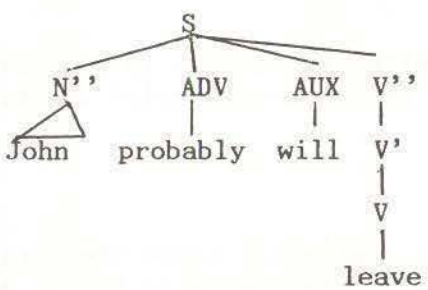
Jackendoff's analysis incorrectly predicts that sentences with S adverbs following the second or third auxiliary should be ungrammatical. But, as Jackendoff (1972) notes, such sentences are not ungrammatical.

16. ?John will have probably been beaten by Bill.
(Jackendoff's example 3.139)

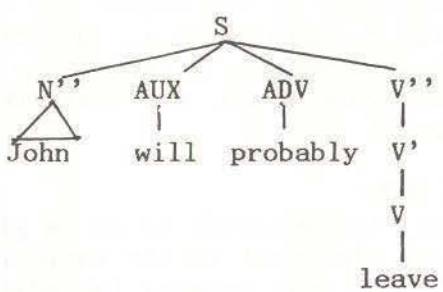
According to Jackendoff's analysis, the surface constituent structures for 17 and 18 would be 19 and 20, respectively.

17. John probably will leave.
18. John will probably leave.

19.

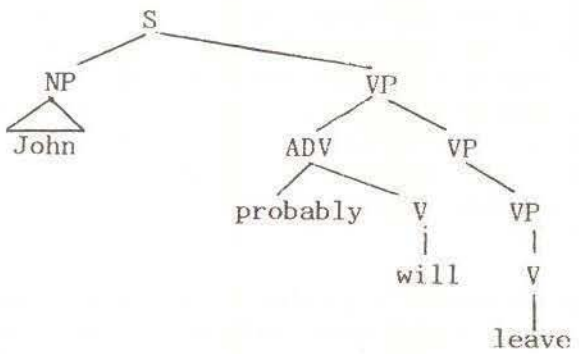


20.

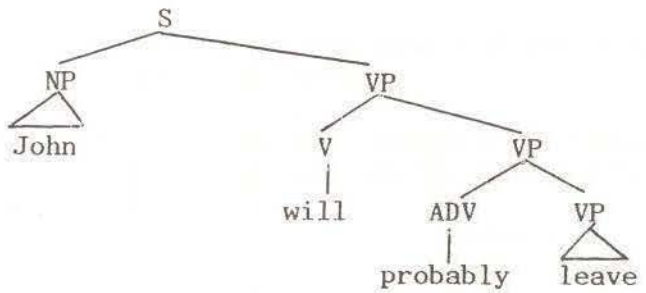


In Section 3, I will argue that the correct constituent structure trees for 17-18 are those in 21 and 22. Since no distinction is made between the matrix VP and embedded VP's in the version of GPSG which I adopt, I use VP, instead of V' or V'', in the trees below.

21.



22.



I will present evidence that sentential adverbs in positions other than clause-initial and clause-final should be accounted for by the phrase-structure rule in 23.

23. VP -> ADV VP

This option was not available to Jackendoff since the phrase-structure rule in 23 does not conform to the rule schema to which, according to the X-bar Convention, all phrase-structure rules must conform. The X-bar Convention requires that one of the daughters in a phrase-structure rule be of the same syntactic category as the mother and one bar level lower than the mother. Thus, the rule in 23 is a counterexample to the X-bar Convention. I will show in chapter 3 that there is ample evidence for the phrase-structure rule in 23, and that, therefore, the X-bar Convention must be rejected.[2]

2.3 Gazdar, Pullum, and Sag (1982)

In Gazdar et al. (1982: 24), the metarule in 24 is given "to handle the facts about sentential adverb placement in the variety of English described by Jackendoff (1972) which only permits the adverb after the first auxiliary verb."

24. <VP -> V VP, F> ==> <VP -> V ADV VP, λP [ADV' (F(P))] >
[+AUX] [-NUL]
[+FIN]

This metarule states that for every rule in the grammar which expands a VP which is marked [+AUXILIARY] and [+FINITE] as V followed by a non-null VP, there will be a rule exactly like this rule except that ADV appears between V and VP. Note that the Head Feature Convention ensures that the V is also [+AUXILIARY] and [+FINITE].

The metarule in 24 is inadequate because it accounts for a very limited range of the positions in which sentential adverbs may occur and also allows for the generation of ungrammatical strings.

Metarule 24 correctly predict that S adverbs occur after the first auxiliary as in 25 and 26.

25. Ed has evidently washed the dishes.
26. Ed will evidently have washed the dishes.

However, as Jackendoff (1972) noted, sentential adverbs may also precede the first auxiliary or main verb. The metarule does not provide for adverbs in these positions. Sentences such as 27-29 are not accounted for by this metarule.

27. Ed obviously has learned French.
28. Ed obviously learned French.
29. Ed obviously will.

The metarule also fails to account for the occurrence of adverbs before the second of two conjoined verbs, as in 30, and for the occurrence of adverbs before the main verbs in sentences in which 'subject-auxiliary inversion' or 'VP fronting' has applied.

30. Ed will catch and probably kill the rabid dog.
31. Will Ed probably kill the rabid dog?
32. ?John said he will definitely pay me and definitely pay me he will.

Gazdar et al. (1982:24) state that "There exists also a less restricted variety in which such adverbs [sentential adverbs] may occur after any auxiliary verb (although the deeper they get in the V', the worse they sound). To handle this variety one needs to delete the [+FIN] specification on the dominant V'." With the [+FIN] specification deleted, the metarule will predict the grammaticality of sentences such as 33.

33. Ed will have evidently washed the dishes.

However, the grammaticality of sentences such as 27-32 is still left unaccounted for.

Gazdar et al. (1982) point out that their metarule predicts the ungrammaticality of strings such as 34 and 35 (their h and i, p. 25) in which the adverb has been stranded.

34. *Kim will obviously.
(with no pause before the adverb)
35. *Kim is obviously
(with no pause before the adverb)

However, this metarule does not predict the ungrammaticality of strings such as 36 and 37.

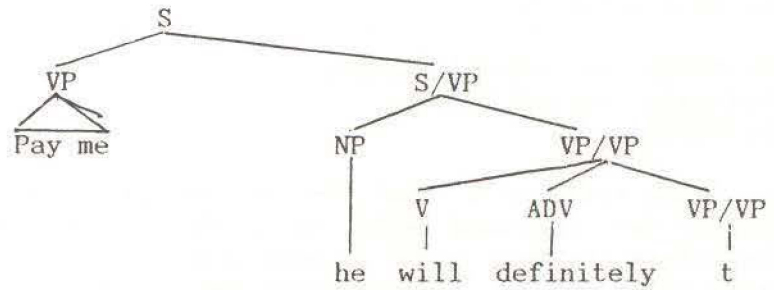
36. *John said he will definitely pay me and pay me he will definitely. (with no pause before the adverb)
37. *I thought John would probably leave and leave he did probably. (with no pause before the adverb)

Metarule 24 in conjunction with the topicalization schema of Gazdar et al. (1982) and the slashing metarule of Gazdar (1981), incorrectly predicts that 36 and 37 are grammatical . The topicalization schema in 38 will allow the rule [S -> VP S/VP].

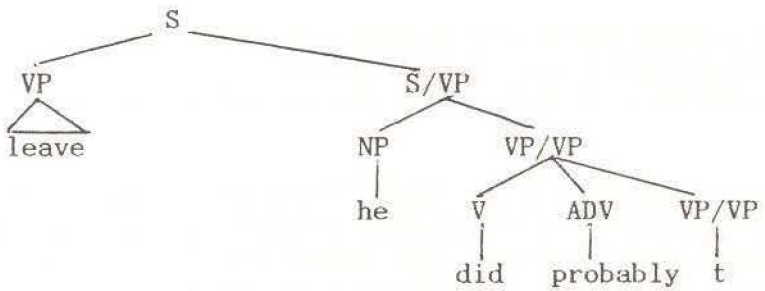
38. $\langle 13, [S \rightarrow a \ S/a], \lambda h_a [(S/a)'](a') \rangle$

The slashing metarule will apply to the rule on the right of the S-Adverb metarule (i.e. [VP -> V ADV VP]) to give the derived rule [VP/VP -> V ADV VP/VP]. These two rules, along with the derived rule [S/VP -> NP VP/VP] will admit the trees in 40 and 41.

39.



40.

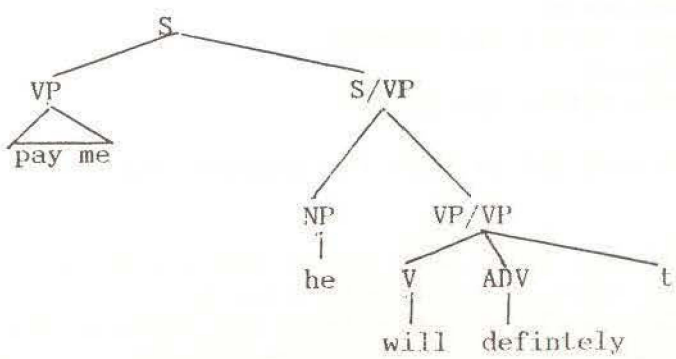


The ungrammatical strings in 36 and 37 will be generated even if the Trace Introduction Metarule (TIM) presented in Sag (1982) is adopted. (The TIM is discussed in more detail in Chapter 6.)

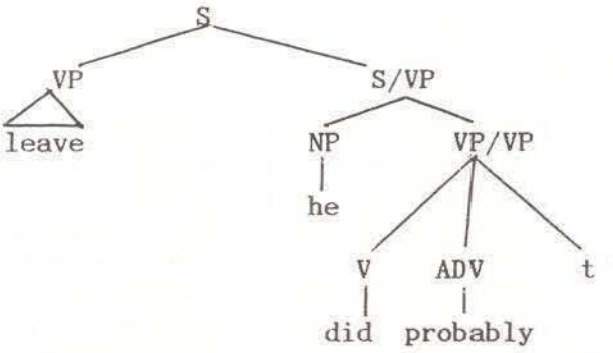
41. Trace Introduction Metarule
[a/B ->...B/B...] ==> [a/B ->...t...]
where a≠B

The TIM requires that the node immediately dominating a trace is of the form a/B where a and B are not identical.
If the TIM is adopted, the trees for 36 and 37 will be 43 and 44.

42.



43.



Presumably these trees are admissible, as well as all other trees representing sentences in which 'VP-fronting' has 'applied', because the features on the two VP's of the VP/VP dominating the trace are not the same and therefore a is not equal to B, as is required. If rules of the form [VP/VP ->...VP/VP...] where the VP's of the dominating VP/VP differ in feature specifications are not allowed as input to TIM, then it would no longer be possible to account for 'VP-fronting'.

Even though 36 and 37 will be generated whether or not the TIM is adopted, the ungrammaticality of such sentences could be accounted for if a surface filter, such as the one proposed by Sag (1978, 1980), is employed.

44. $* \left\{ \begin{matrix} Q \\ ADV \end{matrix} \right\}$ t

This filter rules out strings in which an adverb (or quantifier) immediately precedes an extraction site. In section 6 arguments will be presented against the surface filter in 44, and it will be shown that, given the analysis proposed in chapter 3, no surface filter is necessary.

In the following section, an analysis within the GPSG framework will be presented which accounts for the data in 27-32 and 36-37, not accounted for by the metarule in 24, as well as other data.

FOOTNOTES

1. In 1972 Jackendoff rejected a transportability analysis of VP adverb positions citing as counterevidence cases of strictly subcategorized adverbs, which only occur in postverbal position. In Jackendoff (1977) it is claimed that strictly subcategorized adverbs are dominated by V' rather than V''. Since only adverbs dominated by S or V'' are subject to the transportability convention, these cases no longer represent counterexamples.
2. Radford (1981:104-106) notes other counterexamples to the X-bar Convention.

3. Evaluative and Modal Adverbs

In this section I will deal with two classes of sentential adverbs which fall into Greenbaum's (1969) category of "attitudinal disjuncts": modal adverbs such as probably, possibly, necessarily, and evaluative adverbs such as unfortunately, luckily, preposterously. Greenbaum does not present any data which would indicate that modal adverbs and evaluative adverbs should be distinguished for syntactic purposes. Jackendoff (1972, 1977) does not distinguish these two classes syntactically. Schreiber (1971:84) claims that "a variety of syntactic arguments can be given...that there are indeed two different types here." The only truly syntactic argument which he gives is that modal adverbs occur in questions, whereas evaluative adverbs do not. In section 3.23, however, it is shown that evaluative adverbs may occur in questions, given the appropriate context. Thus, this purported syntactic difference disappears. In the analysis to be presented modal and evaluative adverbs will be given a uniform syntactic treatment and will belong to the same syntactic class. Such a treatment is possible, because modal and evaluative adverbs are of the same semantic type. When in clause-initial and clause-final positions, they are functions from sentence denotations (i.e. denotations of type $\langle s,t\rangle$) to sentence denotations--they are of the type $\langle \langle s,t\rangle,\langle s,t\rangle\rangle$. When in other positions, it will be claimed, they are functions from VP denotations $\langle\langle s,\langle\langle s,\langle e,t\rangle\rangle,t\rangle\rangle,\langle s,t\rangle\rangle$ to VP denotations--they are of type $\langle\langle\langle s,\langle\langle s,\langle e,t\rangle\rangle,t\rangle\rangle,\langle s,t\rangle\rangle,\ \langle\langle s,\langle\langle s,\langle e,t\rangle\rangle,t\rangle\rangle,\langle s,t\rangle\rangle\rangle$.[1]

It is useful to consider the adverbs which Schreiber (1971:88) assigns to each category.

1. Modal adverbs: allegedly, certainly, conceivably, evidently, possibly, undoubtedly, unquestionably, clearly, obviously, apparently.
2. Evaluative adverbs: unfortunately, predictably, regrettably, astonishingly, incredibly, interestingly, ironically, luckily, naturally, oddly, predictably, strangely, surprisingly, unbelievably, understandably, unluckily.

What distinguishes these two classes from one another semantically is that the evaluative adverbs are factive, whereas the modal adverbs are not (i.e. Unfortunately, John left presupposes, and perhaps entails, that John left, but Possibly, John left does not).

The analysis to be presented accounts for the occurrence of modal and evaluative adverbs in sentences in which the adverb has scope over the rest of the sentence. I will not deal with the positioning of adverbs in sentences such as 3-5 in which the adverb does not have scope over the rest of the sentence. In 3 probably has scope only over the prepositional phrase in Westerville. In 4 probably has scope only over the verb phrase sing a maudlin song. In 5 the adverb has scope

only over the NP Sharon.

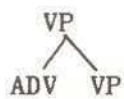
3. We plan to buy a house, probably in Westerville.
4. John will do something for amateur night, probably sing a maudlin song.
5. I gave the book to one of my students, probably Sharon.

I have set aside such sentences from consideration, because I believe that the syntactic analysis of these adverbs will be independent of the syntactic treatment of adverbs in other positions, and thus not immediately relevant to the analyses to be given.

3.1 Evaluative and modal adverbs in positions other than clause-initial and clause-final

In this section it will be argued that modal and evaluative adverbs, when in positions other than clause-initial and clause-final, and when not requiring the intonation pattern required by parentheticals, appear in the configuration in 6.

6.



I will sometimes refer to an adverb in the configuration in 6 as being 'Chomsky-adjoined' to the VP, meaning only that the adverb occurs in this configuration, not that it is actually placed there by a transformation. The lower VP in 6 may dominate either a main or an auxiliary verb.

I will assume that sentences in which the adverb is both preceded and followed by a pause are structurally distinct from sentences in which the adverb is not preceded or followed by a pause. The sentences in 7 and 8, for example, will be assigned distinct structures.

7. John will unfortunately leave.
8. John will, unfortunately, leave.

The adverb in 7 will appear in the configuration in 6, but in 8 it will not. I will assume that the sentence in 8 will have a structure identical to that which a sentence such as 9 has, whatever that may be.

9. John will, as you know, leave.

I am assuming the treatment of auxiliaries given in Gazdar et al. (1982). In this treatment auxiliaries are introduced as daughters of VP by the finite rule schema in 10. The use of features ensures that co-occurence restrictions involving auxiliaries are met. Note that the infinitive marker to is also considered to be a verb.[2]

10. <n, [VP -> V VP], P[V'(VP'(P))]>
[a] [B]
[+AUX]
where values for n, a, and B are given by Table 1.

[n]	a	B	V[n] MEMBERSHIP
[2]	+FIN	+BSE	can, may, must, will etc.
[3]	+FIN	+BSE,-AUX	do
[4]	+ASP	+PSP	have
[5]	+ASP,+COP	+PRP	be
[6]	+COP	+PAS	be
[7]	+INF	+BSE	to
[8]	+FIN,+COP	+INF	is[+COP],ought[-COP]
[9]	+COP	+PRD	be

TABLE 1

My claim is that modal and evaluative adverbs in all of the positions below are 'Chomsky-adjoined' to the following VP.

i. Before a finite main verb:
11. Mark probably left.
12. Mark unfortunately left.

ii. Before a finite auxiliary verb:
13. John probably will leave.
14. John unfortunately will leave.
15. John probably will.

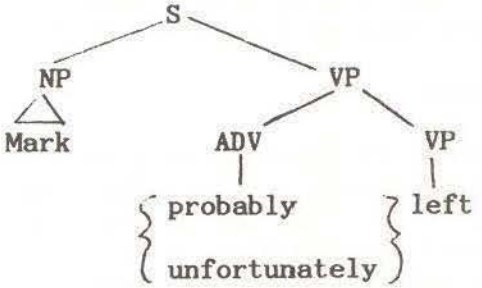
iii. Between a finite auxiliary verb and the main verb:
16. John has probably left.
17. John has unfortunately left.

iv. Between a nonfinite auxiliary verb and the main verb:
18. Ed will have probably washed the dishes by now.
19. Ed will have fortunately washed the dishes by now.

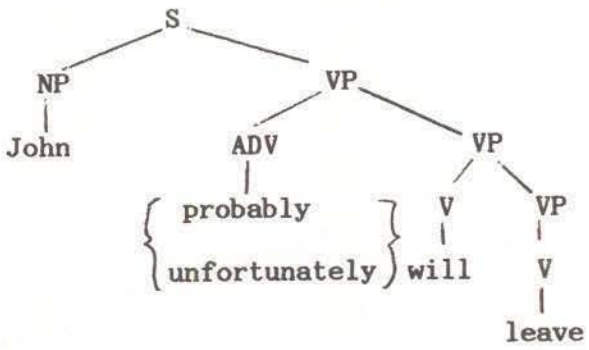
v. Between any two auxiliary verbs.
20. Ed will probably have washed the dishes by now.
21. Ed will fortunately have washed the dishes by now.

The trees for sentences 11-21 are given below.

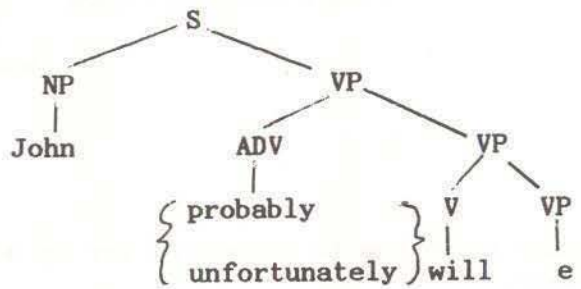
22.



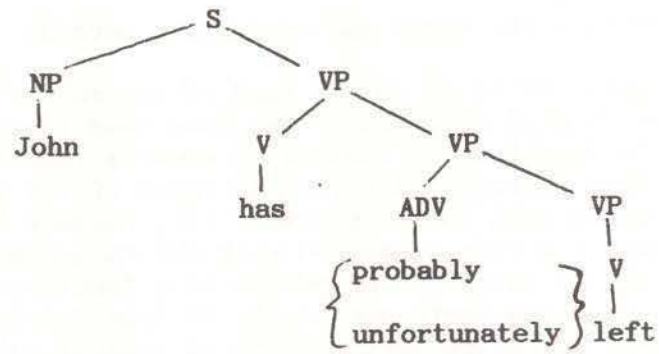
23.


S
NP
VP
John
ADV
VP
probably
unfortunately
V
VP
will
V
leave

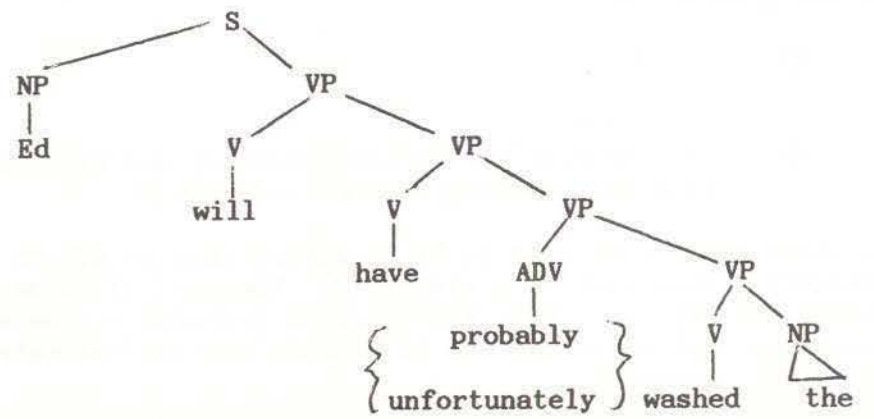
24.


S
NP
VP
John
ADV
VP
probably
unfortunately
V
VP
will
e

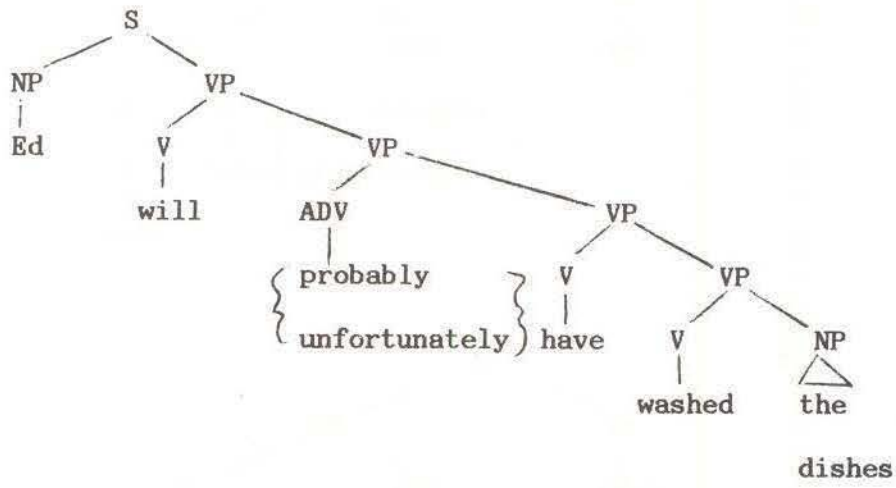
25.


S
NP
VP
John
V
VP
has
ADV
VP
probably
unfortunately
V
left

26.


S
NP
VP
Ed
V
VP
will
V
VP
have
ADV
VP
probably
unfortunately
V
NP
washed
the
dishes

27.



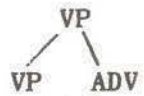
All of the sentences above can be accounted for by adding the ID rule in 28 to the grammar.

28. <1, [VP -> ADV, VP], ADV'(VP')>
[1]
where ADV(1) = the modal and evaluative adverbs

Note that the daughter VP in 28 is the head of the mother VP. Therefore, the daughter VP will have all of the same head features as its mother, because the Head Feature Convention ensures that the head of a phrase will have head features identical to those of its mother node. The ADV will be marked with the rule number (1), because the lexical category introduced by a rule is marked with the rule number. In the semantic rule in 28 the value of the adverb is a function which takes VP type denotations as arguments and yields VP type denotations.

It is necessary to ensure that adverbs of lexical category 1 may not follow their sister VP's. Otherwise, the subtree in 29 will be generated and it will be incorrectly predicted that sentences such as 30 are grammatical.

29.



30. *Patrick went to the bank probably and withdrew money from from his checking account definitely.

We could add the LP rule in 31 to ensure that an adverb of the lexical category 1 must precede a sister VP. However, if we assume, as do Gazdar and Pullum (1982), that English includes a general LP rule requiring lexical categories to precede non-lexical categories, then 30 is not necessary.

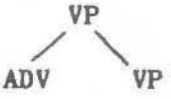
31. ADV <VP
[1]

Given the ID rule in 28 and the LP rule in 31 the basic phrase-structure rule in 32 below will be part of the grammar, but the rule in 33 will not.

32. S-Adverb basic rule:
<1, [VP -> ADV VP], ADV' (VP')>
where ADV(1)= the evaluative and modal adverbs
33. <1, [VP -> VP ADV], ADV' (VP')>

I will label 32 the S-Adverb basic rule, since evaluative and modal adverbs have traditionally been known as sentential adverbs, but I do not mean to imply that every adverb which has been labeled a sentential adverb should be introduced by a rule like that in 28.[3]

In the following section I will show that once the rule in 32 is adopted, interaction with a number of independently-motivated rules of GPSG accounts for a wide range of data. First, however, I will use some of this data as basically theory-neutral evidence for the configuration in 6, repeated below.

34.


The examples in 35-37 provide evidence for the higher VP node in 34.

35. John said he would definitely pay me and definitely pay me, he will. [VP Preposing]
36. Two plus two will necessarily equal four and one plus three will, too. [VP Deletion]
37. John probably will swing and possibly will hit the ball. [VP Conjunction]

If 35-37 are indeed examples of VP Preposing, VP Deletion, and VP Conjunction, as they certainly appear to be, then the adverb and following verb phrase must be dominated by VP.

Examples such as 38 provide evidence for the lower VP node in 34.

38. Rhonda has probably been to Dinosaur Park and Jimmy definitely has.

Because of the presence of definitely in the second conjunct, the only interpretation for this sentence is one in which only been to Dinosaur Park has been 'deleted', and not probably been to Dinosaur Park. The semantic rule associated with VP Deletion ensures that the value of a previous VP is eventually plugged in to the translation of the right conjunct. Since the value of been to Dinosaur Park is plugged in to the translation of the right conjunct in 38, been to Dinosaur Park in the left conjunct must be a VP. A semantic analysis of VP Deletion must allow either the value of the lower VP or the value of the higher VP to be plugged in to the translation of the right conjunct. In 36, the most natural reading is one in which the value of the higher VP is

plugged in. In 38, the reading in which the value of the lower VP has been plugged in is forced by the presence of definitely.

3.2 Evidence from rule interactions

In this section I will show that once the S-Adverb basic rule in 32 is adopted, interaction with a number of independently-motivated rules of GPSG account for a wide range of data.

3.21 Coordination

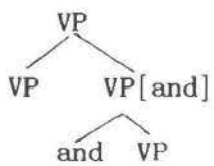
Gazdar (1981:158) proposes the following rule schemata to account for constituent coordination.

39. $\langle 2, [a \rightarrow a_1 \ldots, a_n], B'(a_1', \ldots, a_n') \rangle$
$[B]$
where B $\in$ [and, or] and a is any syntactic category

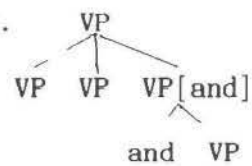
40. $\langle 3, [a \rightarrow B\ a], a' \rangle$
$[B]$
where B $C-$ [and, or] and a is any syntactic category

If a is VP and B is and, then the schemata in 39 and 40 will produce structures such as 41 and 42.

41.



42.

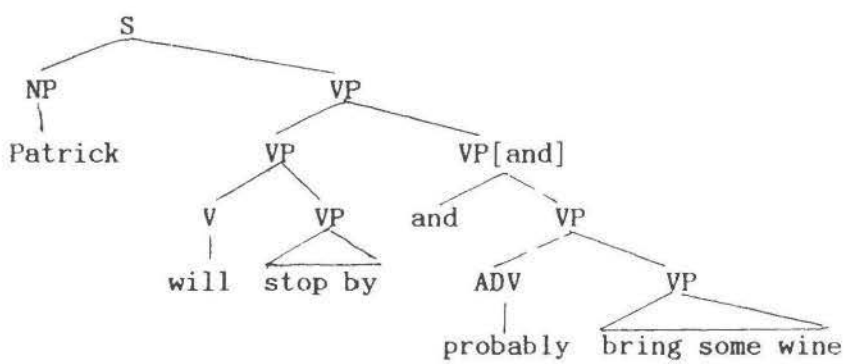


These rule schemata together with the S-Adverb basic rule proposed in section 3.1 predict the grammaticality of sentences such as 43-46 in which adverbs precede verb phrase conjuncts.

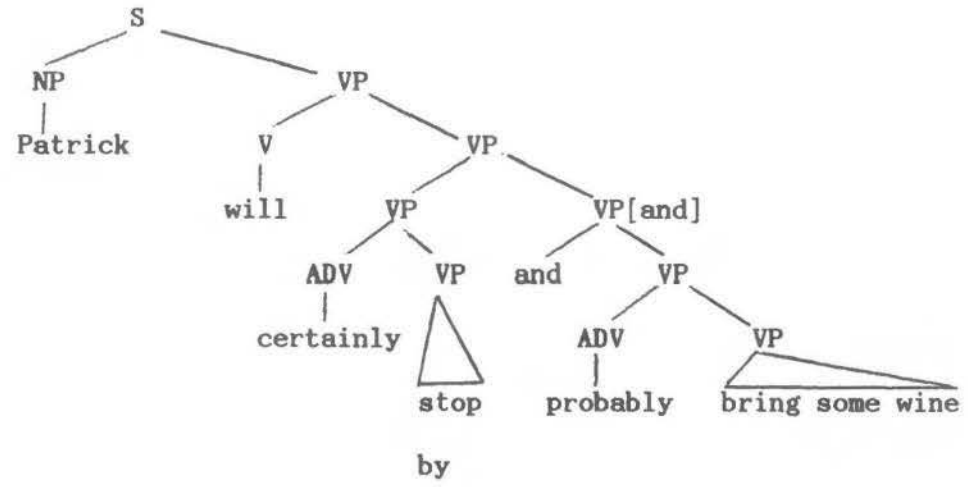
43. Patrick will stop by and probably bring some wine.
44. Patrick will certainly stop by and probably bring some wine.
45. Patrick studied, but probably flunked the test anyway.
46. Patrick probably works hard and definitely enjoys his work.

These sentences will be assigned the following trees:

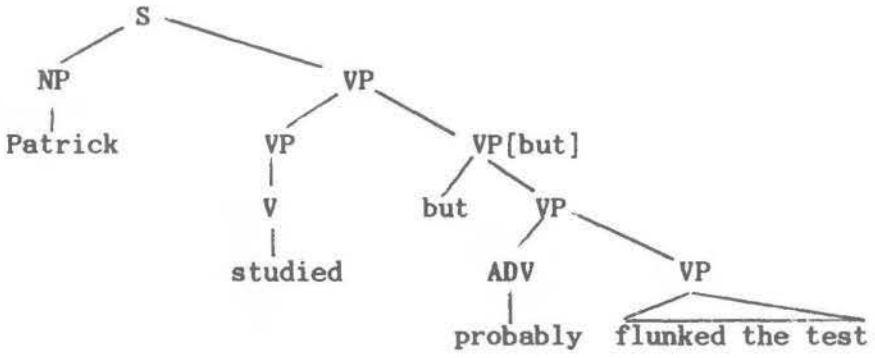
47.



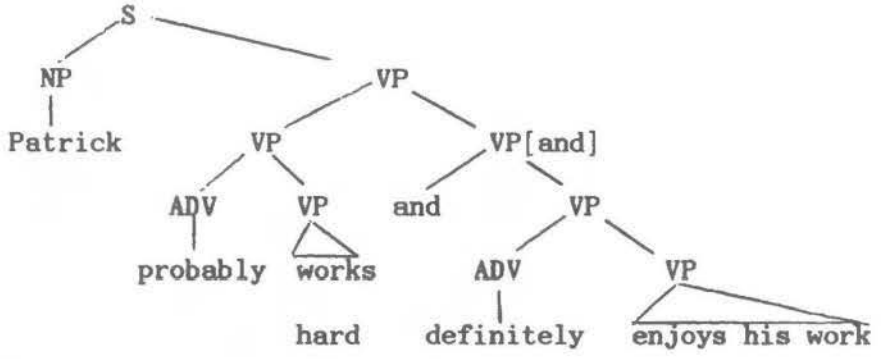
48.



49.



50.



The grammaticality of sentences such as 51-54, with evaluative adverbs, is, of course, also predicted.

51. Patrick will stop by and unfortunately stay for dinner.
52. Patrick will fortunately stop by, but unfortunately stay for dinner.
53. Patrick stopped by and unfortunately stayed for dinner.
54. Patrick fortunately stopped by, but unfortunately stayed for dinner.

Examples such as 43-46 and 51-54 are problematic for Jackendoff's analysis. They are obvious counterexamples to the claim that S-adverbs always occur as daughters of S at the surface structure level. Once Conjunction Reduction is given up, such examples are also counter-examples to the claim that S-adverbs are always daughters of S at the deep structure level.

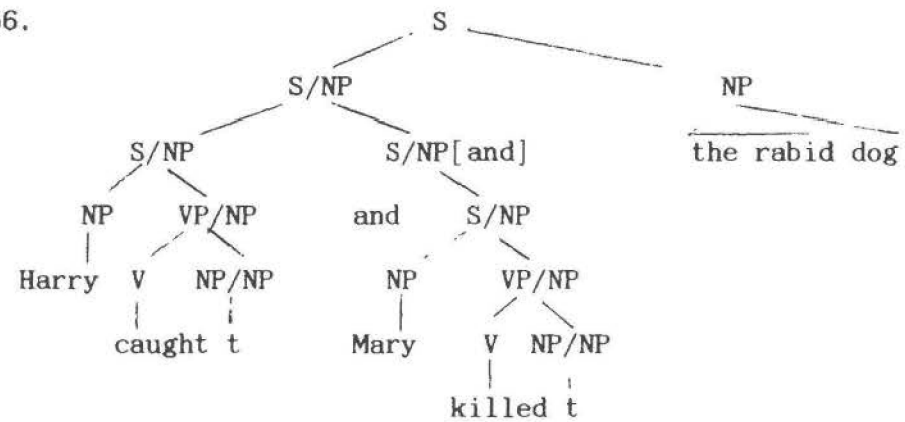
3.22 Right Node Raising

The next set of examples include sentences which have traditionally been described as having undergone the transformation of right node raising. In Gazdar (1981) right node raising structures are accounted for by the rightward displacement schema in 55.

55. $\langle 9, [a \rightarrow a/B\ B], \lambda h_B[(a/B)'](B')\rangle$
where a ranges over clausal categories and B can be any phrasal or clausal category.

The rightward displacement schema, the coordination schemata, and the slashing metarule interact to produce structures such as that in Gazdar's example in 56.

56.



The rightward displacement schema permits the rule [S -> S/NP NP].

The coordination schemata permit the rules [S/NP -> S/NP S/NP] and [S/NP -> and S/NP].

The slashing metarule permits the rules [S/NP -> NP VP/NP] and [VP/NP -> V NP/NP].

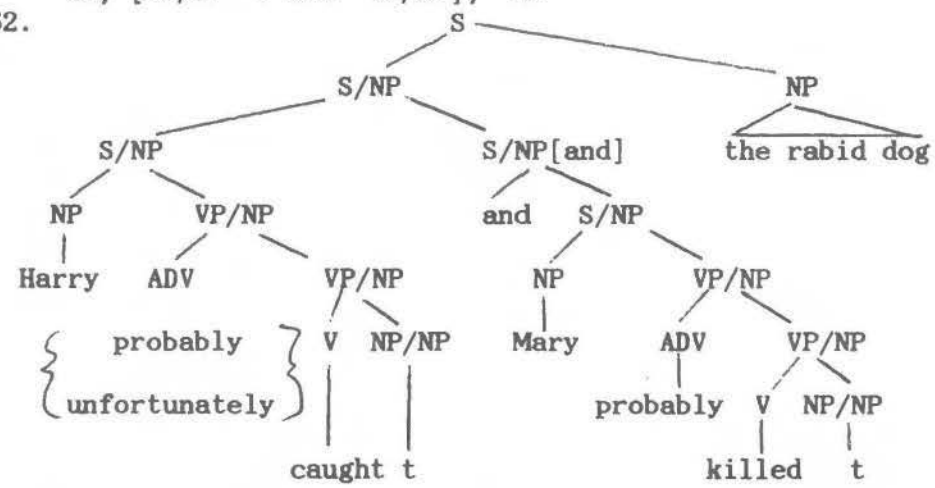
With the addition of the S-Adverb basic rule to the grammar, the grammaticality of right node raising sentences such as the following in which an S-adverb precedes the verb(s) is predicted.

57. Harry probably caught and Mary certainly killed the rabid dog.
58. Harry caught and Mary probably killed the rabid dog.
59. Harry caught and Mary unfortunately killed the rabid dog.
60. Harry fortunately caught and Mary fortunately killed the rabid dog.

The slashing metarule applies to the S-Adverb basic rule, as in 61, and allows the derived rule on the righthand side of the arrow. This derived rule allows the adverb to precede a verb which has had its object raised, as shown in 62.

61. <1, [VP -> ADV VP], F> ==>
<1, [VP/NP -> ADV VP/NP], F>

62.

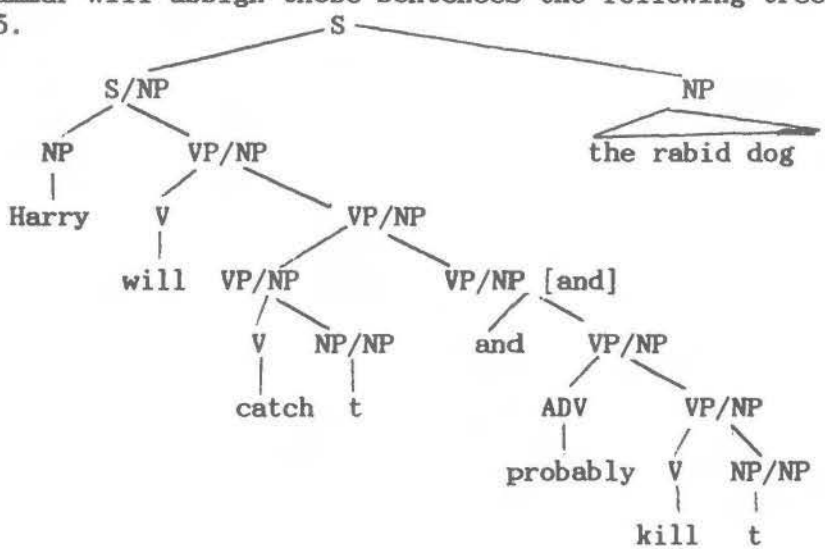


The rightward displacement schema, the coordination schemata, the slashing metarule, and the derived rule in 61, together, predict that sentences such as the following are grammatical.

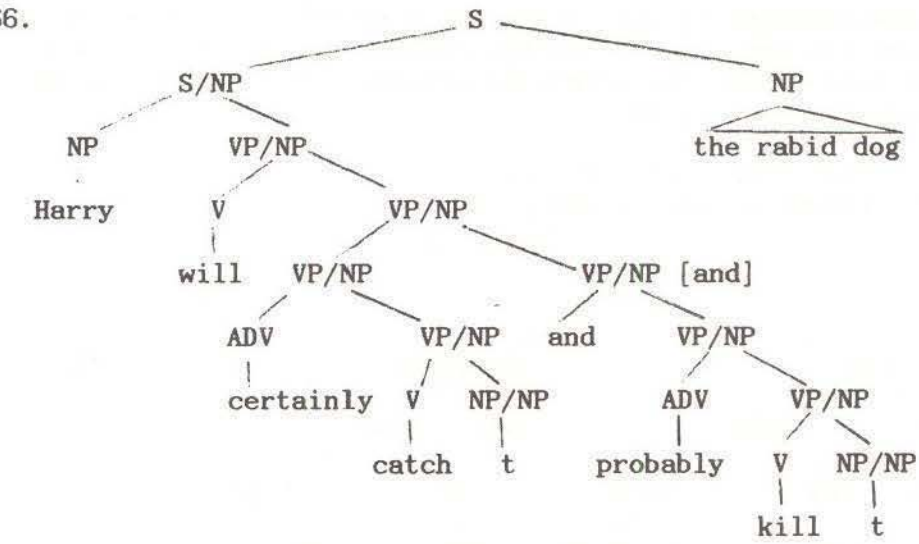
63. Harry will catch and probably kill the rabid dog.
64. Harry will certainly catch and probably kill the rabid dog.

The grammar will assign these sentences the following trees.

65.



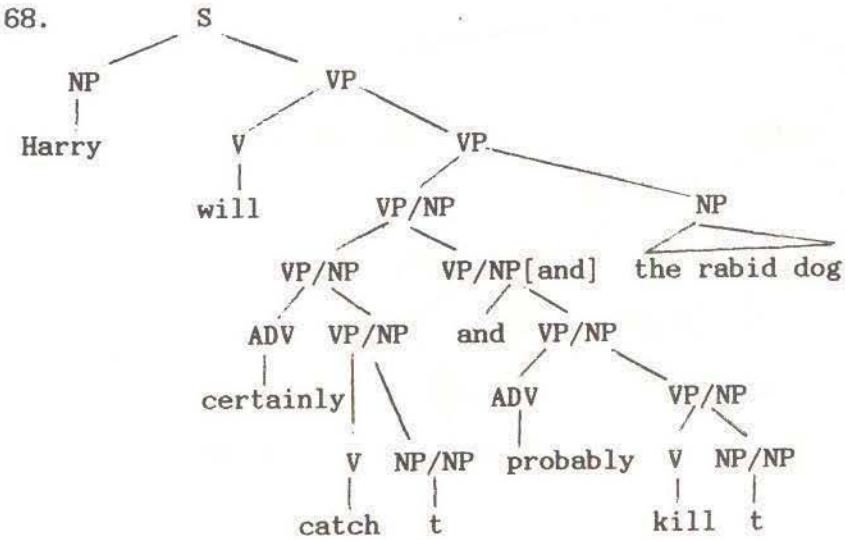
66.



The trees in 65 and 66 are right node raising structures, but sentences such as 63 and 64 can be produced without the intonation pattern characteristic of right node raising. Thus, it seems necessary to provide trees for such sentences which do not have a right node raising configuration. I will argue that the grammar should include the rule of "minor right node raising" in 67.[4] Once this rule is adopted, sentences such as 63 and 64 will be assigned two distinct tree structures, one like those in 65 and 66 and another like that in 68.

67. <57, [VP -> VP/NP NP], VP/NP'(NP')>

68.

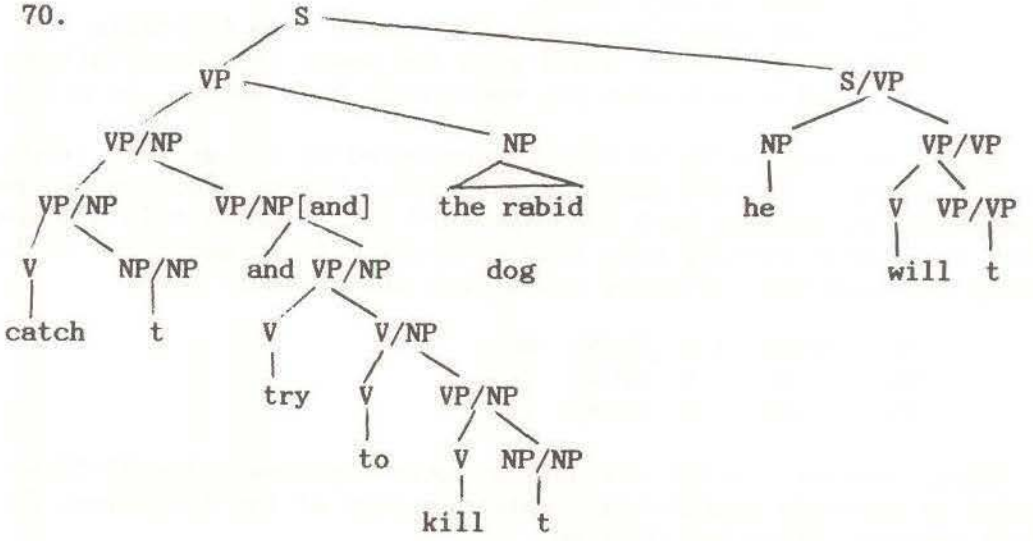


Inclusion of the rule [VP -> VP/NP NP] in the grammar can be motivated by considering sentences such as 69.

69. Ed said he would catch and try to kill the rabid dog and catch and try to kill the rabid dog he will.[5]

If the second conjunct of sentence 69 is to be treated as an example of 'VP-fronting', then the grammar must provide a structural description of catch and try to kill the rabid dog in which the entire phrase is a VP. The minor right node raising rule provides for such a description, as shown in 70.

70.

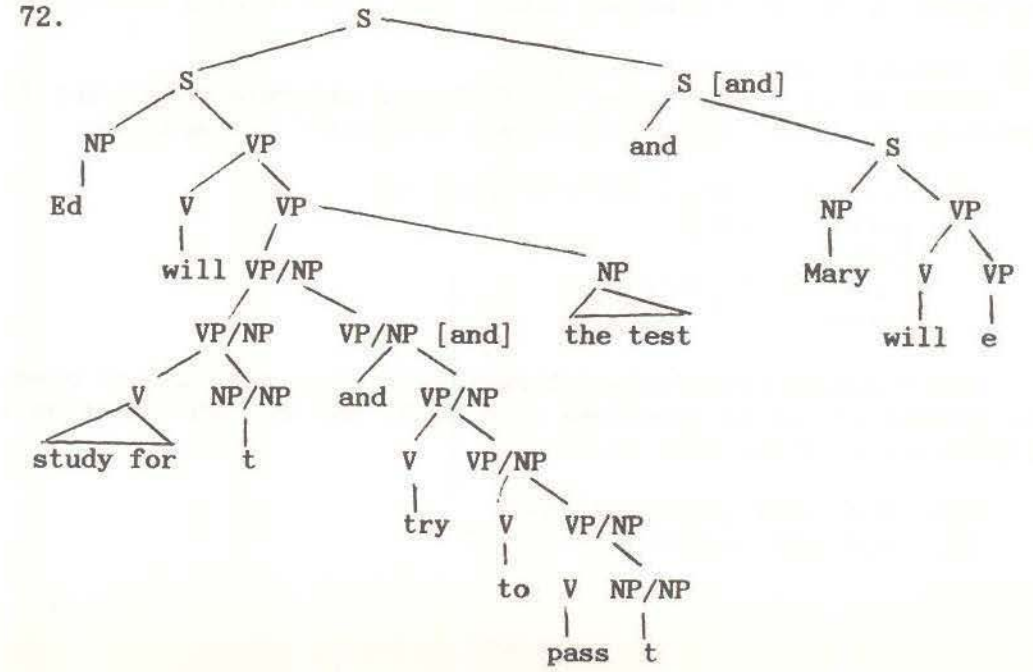


Sentences such as 71 provide further evidence for minor right node raising.

71. Ed will study for and try to pass the test and Mary will too.

If the second conjunct in sentence 71 is to be treated as an example of 'VP Deletion', then the grammar must allow study for and try to pass the test to be a VP. The rule [VP -> VP/NP NP] will provide a representation in which it is a VP, as in 72.

72.



It must be noted that inclusion of the rule [VP -> VP/NP NP] will allow for the generation of the following ungrammatical sentences:

73. *John gave a vase the woman who he's dating.
74. *John persuaded that Harold left the woman who's standing over there.
75. *John expected to win the runner from Australia.
76. *John said he would give the woman he's been dating a vase and give a vase the woman he's been dating, he will.

Sentences such as 73-75 will be generated by the grammar anyway because of the rightward displacement rule. One solution which comes to mind for preventing both the rightward displacement rule and the minor right node raising rule from generating these sentences is to simply disallow the following rules from the grammar:

77. VP/NP -> V NP/NP NP
78. VP/NP -> V NP/NP S
79. VP/NP -> V NP/NP VP

This, however, is not a viable solution because rules 77-79 are needed to generate grammatical sentences such as the following, in which topicalization has applied.

80. Janet, John gave a vase.
81. Kim, John persuaded that Fido runs.
82. Jimmy Carter, John wanted to win.

Unfortunately, I do not have a solution to offer at this time. It should be noted, however, that examples such as 73-75 have been problematic for all previous analyses of right node raising, and that it is very likely that an account of the ungrammaticality of these sentences when assigned right node raising structures will also account for their ungrammaticality when assigned minor right node raising structures.

3.23 Subject-Auxiliary Inversion

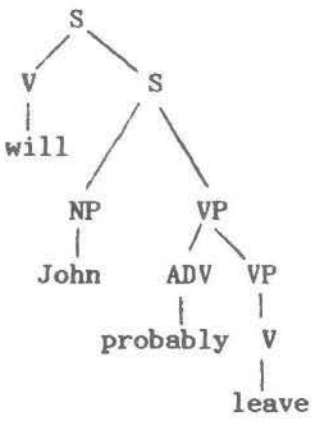
Gazdar et al. (1982) give the following metarule to account for sentences in which 'Subject-Auxiliary Inversion' has applied.

83. $\langle$[VP -> V VP], $\hat{\lambda}$P[V'(^V'(P))] ==>
[+FIN] [α]
[+AUX]
$\langle$[S -> V S], V'(^S')$\rangle$
[+INV] [α]

The S-Adverb basic rule interacts with this metarule to predict the grammaticality of questions such as 84 and 85. The question in 84 will be assigned the tree in 86.

84. Will John probably leave?
85. Did John fortunately leave?

86.



Both Jackendoff (1972) and Bellert (1977) have argued that questions such as 84 with modal adverbs are syntactically well-formed, but semantically or pragmatically odd.

Jackendoff (1972:84) notes that "many S adverbs do not feel comfortable in questions" and 'stars' 87 (his 3.160) to indicate that it is unacceptable.

87. *Did Frank probably beat all his opponents?

He argues that "purely syntactic approaches to the unacceptability of sentences such as 87 miss the point" and that "a more semantically based analysis is called for, in which there is a reason for these facts" (p. 84).

Bellert (1977:344) states that "Modal sentential adverbs are predicates of the truth: they qualify the truth of the proposition expressed in the same sentence, and they do not qualify it negatively. Neither do they occur in questions" and cites 88 (her 21).

88. *{Has / Will} John {probably / certainly / evidently} come?

Such questions are unacceptable, according to Bellert, because "we do not ask questions and at the same time evaluate the truth, or degree of truth, of the proposition that is being questioned" (p. 344).

The explanation offered by Bellert is supported by the observation that, in contexts in which these constraints do not, in general, hold, questions such as those in 87 and 88 are acceptable. A context in which we expect questions to be asked which "at the same time evaluate the truth, or degree of truth, of the proposition that is being questioned" and which "assert a proposition in one and the same sentence" is the courtroom context. Questions such as 89-90 are certainly acceptable in a courtroom setting.

89. In your opinion, has the defendent possibly perjured himself?
90. In your opinion, did John Jones probably commit suicide?

It is clear from such examples that the unacceptability of sentences such as 87 and those in 88 should be accounted for by constraints such as those offered by Bellert, rather than by syntactic constraints. A similar constraint is obviously responsible for the

unacceptability of questions such as 85: we do not question a proposition and at the same time comment on the proposition we are questioning. This constraint is violated in a courtroom setting, perhaps because lawyers are often obviously assuming the truth of a proposition while at the same time asking for a witness to evaluate the truth of the proposition. So, in 91, for example, the questioner is commenting on the proposition, while at the same time asking that the respondent evaluate its truth.

91. Did this woman unfortunately get involved in a life of crime?

I will conclude that the questions in 84-85 and 87-88 are syntactically well-formed and that the S-Adverb basic rule and 'Subject-Auxiliary Inversion' metarule interact to yield correct predictions.

3.24 VP Fronting

Gazdar et al. (1982) account for 'VP-fronting' by the more general rule of Topicalization, repeated below. They claim that "the phenomenon commonly referred to as 'VP-fronting' is simply a special case of topicalization and can therefore be subsumed under schema 13 [92 below] by allowing a to range over frontable V' [VP] types"(p. 18).

92. $\langle 13, [S \rightarrow a\ S/a], \lambda h_a[(S/a)'](a')\rangle$
when a = VP, then a is to be [-FIN, INF, ASP]

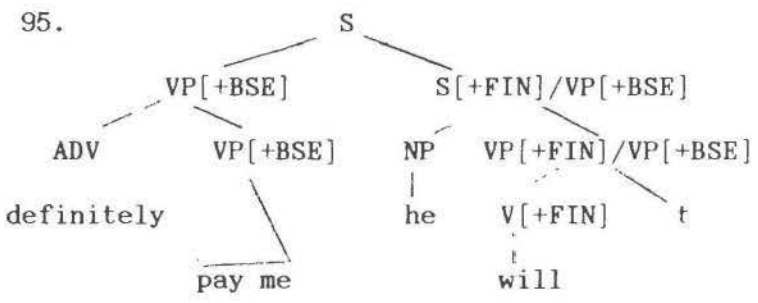
The S-Adverb basic rule and the topicalization schema in 92 correctly predict that sentences such as 93 and 94 are grammatical.

93. ?John said he will definitely pay me and definitely pay me he will.
94. ??I thought John would probably leave to avoid seeing his mother and probably leave he did.

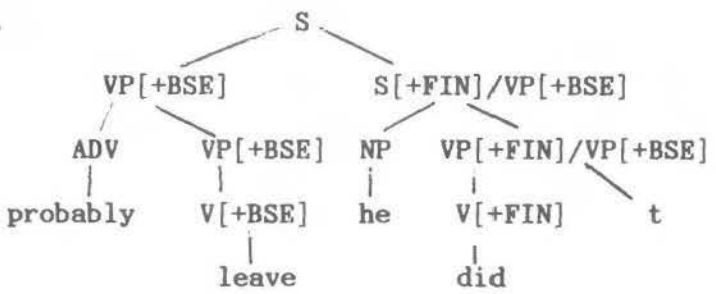
While sentences such as 93 and 94 may not be fully acceptable, they are certainly not ungrammatical. It seems likely that semantic and/or pragmatic constraints on VP Topicalization, similar to those discussed by Prince and Prince (1980) for NP Topicalization, may be responsible for the oddness of such sentences.

These sentences will be assigned the structures in 95 and 96. The topicalization schema allows the rule [S -> VP S/VP] and the S Adverb basic rule allows the rule [VP -> ADV VP].

95.



96.



It is also predicted that sentences such as 97 are grammatical, since evaluative adverbs are also introduced by the S-Adverb basic rule.

97. Bob knew Bill would unfortunately flunk the test, and unfortunately flunk the test, he will.

However, sentences such as 97 are much worse than sentences such as 93 and 94. But the differences in acceptability are expected, given the differences in acceptability between 98 and 99.

98. Bob knew Bill would probably flunk the test.
99. Bob knew Bill would unfortunately flunk the test.

Sentence 98 is fine, but 99 is questionable. The oddness of sentences such as 99 obviously has to do with the function of evaluative adverbs: they are expressions of the speaker's attitude. Parentheticals such as I think also express the speaker's attitude. Although parentheticals which express the speaker's attitude occur freely between the auxiliary and VP in main clauses, as in 100, they do not occur in this position in embedded clauses.

100. Bill will, I think, flunk the test.
101. Bob knew Bill will, I think, flunk the test.

The oddness of sentences such as 99 should be explained in part by the same pragmatic or semantic constraint which explains the oddness of sentences such as 101.

Given the rule of Topicalization and the derived rule in 102, it would seem that ungrammatical strings such as that in 103, in which the adverb is stranded, will result.

102. VP/VP -> ADV VP/VP
103. *John said he would definitely pay me, and pay me he will definitely. (with no pause before the adverb)

However, once the Trace Introduction Metarule is adopted, strings such as that in 103 will not be produced. I will not explain here why this is the case, since a detailed explanation is given in section 6.

3.25 VP Deletion

Gazdar et al. (1982) give the following metarule to account for 'VP Deletion':

104. VPD: < [VP -> V VP], F > --> <[VP -> V VP], F>
[+AUX] [+NUL]
[-PRP]
[GER]

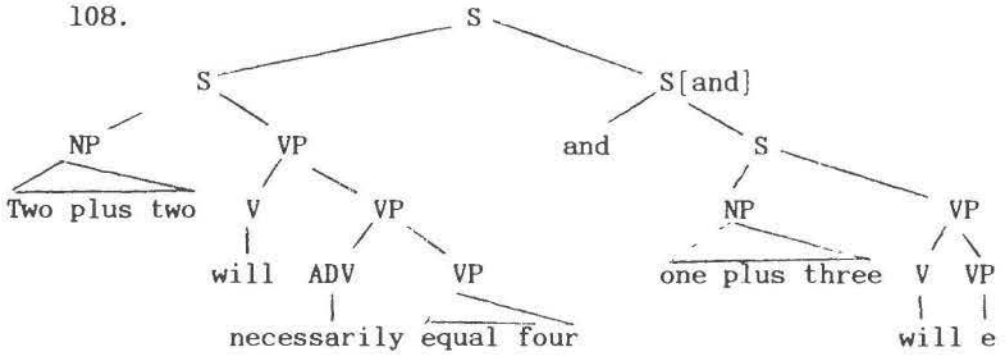
The metarule in 104 "takes any V' [VP] [+AUX, -PRP, GER] rule which expands as V followed by V' [VP], and simply adds the feature +NUL to the complement V' [VP]" (p. 606). The rule in 105 introduces e, which represents the empty string.

105. <[VP -> e, v >
[+NUL]
where v is a contextually bound variable ranging over VP denotations.

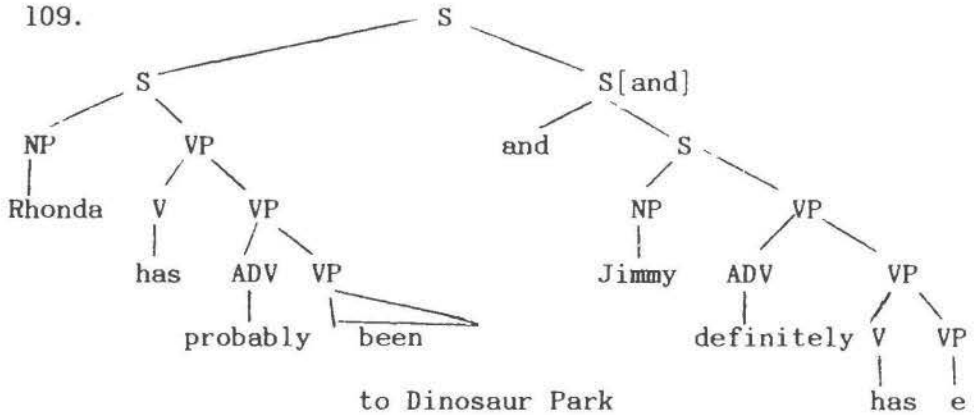
Given this analysis of VP Deletion and the S-Adverb basic rule, sentences such as 106 and 107 will be assigned the trees in 108 and 109, respectively.

106. Two plus two will necessarily equal four and one plus three will, too.
107. Rhonda has probably been to Dinosaur Park and Jimmy definitely has.

108.



109.

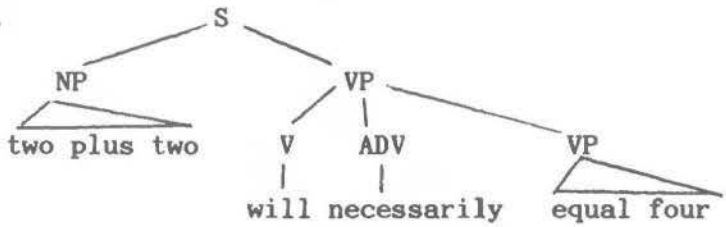


The most natural reading of 106 is the one in which the contextually bound variable v takes as its value the denotation of the VP

necessarily equal four. In 107, the reading in which v takes as its value the denotation of the VP been to Dinosaur Park is forced by the presence of definitely in the second conjunct.

It should be noted that given the metarule analysis of sentential adverbs of Gazdar et al. (1982) it would be difficult to explain the interpretation of 106 in which what has been 'deleted' semantically in the second conjunct is necessarily equal four, since, according to this analysis, necessarily equal four would not be a constituent. Under their analysis, 106 would be assigned the tree in 110.

110.



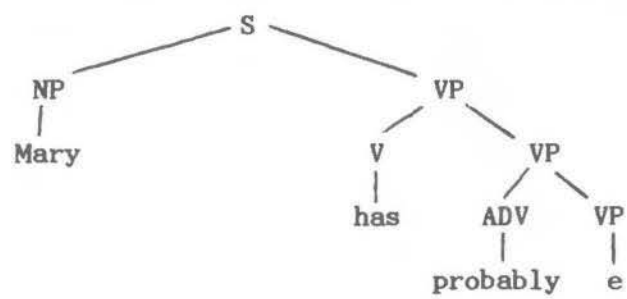
Jackendoff's (1972, 1977) analysis would encounter the same problem, because his analysis also claims that necessarily equal four is not a constituent. Thus, sentences such as 106 are problematic for the analyses of Jackendoff (1972, 1977) and Gazdar et al. (1982), but immediately accounted for by the S-Adverb basic rule, given the 'VP Deletion' metarule of Gazdar et al. (1982).

Given the VP Deletion analysis and the S-Adverb basic rule, it would seem that ungrammatical strings such as that in 111 would be generated.

111. *John has probably gone to Cleveland and Mary has probably, too.
(with no pause between the auxiliary verb and adverb)

The second conjunct would apparently be assigned a structure as in 112.

112.



The VP Deletion metarule allows the rule [VP -> V VP] . The S-Adverb
[+null]
basic rule allows [VP -> ADV VP]. The Head Feature Convention,
[+null]
assuming as do Gazdar et al. (1982) that null is a head feature, ensures that the lower VP is also [+null]. We could account for the ungrammaticality of sentences such as 111 simply by revising the S-Adverb basic rule as in 113, specifying that the dominating VP must be [-null].

113. VP -> ADV VP
[-null]

However, in section 6, I will argue that it is not necessary to specify that the dominating VP is [-null] in order to account for the ungrammaticality of sentences such as 111.

3.2 Evaluative and modal adverbs in clause-initial and clause-final positions

Evaluative and modal adverbs occur in clause-initial and clause final positions with comma intonation, as the following examples show.[6]

114. Unfortunately, John has been in an accident.
115. Obviously, he was driving while he was drunk.
116. He was not seriously injured, fortunately.
117. He will be released soon, probably.

They also occur in initial and final positions in embedded clauses.

118. Mike knows that unfortunately John has been in an accident.
119. The legend that Milton was an unpopular poet has lived so long that probably it will never be destroyed.[7]

The adverb must follow the complementizer or it will not be interpreted as part of the embedded clause. The only interpretation of 120 for example, is one in which unfortunately has scope over the root sentence and not just the embedded clause. Such sentences require parenthetical intonation.

120. Mike knows, unfortunately, that John has been in an accident.

Gazdar et al. (1982) mark that-clauses with the feature [+Complementizer]. That-clauses are introduced by the basic rule in 121.

121. <6, [S -> that S], S'>
[+C] [-C]

The ID rule in 122 accounts for evaluative and modal adverbs in clause-initial and clause-final positions. Marking S as [-Complementizer], ensures that the adverb will not be part of the embedded clause if it precedes the complementizer.

122. <2, [S -> ADV, S], ADV'(S')>
[-C]
where ADV = the evaluative and modal adverbs
[2]

We now have two lexical classes consisting of modal and evaluative adverbs. Adverbs of lexical class 1 are introduced by the ID rule <1, [VP -> ADV VP], ADV'(VP)> and adverbs of class 2 are introduced by the ID rule in 122. It is, of course, desirable to relate adverbs of lexical class 2, which are of type <<s,t>,<s,t>>, with their doublets in lexical class 1. The lexical rule in 123 will accomplish this (cf. Dowty (1978) for an explanation of lexical rules in Montague Grammar).

123. If a θ ADV [1], then $F_i(\alpha) \in$ ADV [2], where $F_i(\alpha) = \alpha$.
Translation: $\lambda\mathcal{P}\lambda P[a'(\mathcal{P}(P))]$, where $\mathcal{P}$ is a variable over VP type denotations and P is a variable over NP type denotation.

This rule states that if there is an adverb which is of lexical class 1, then there is a corresponding adverb of lexical class 2 which has the same form, and which is translated as $\lambda\mathcal{P}\lambda P[a(\mathcal{P}(P))]$.

If 122 is revised so that the lexical class 2 includes other adverbs which occur clause-initially and finally, then there will not be a one-to-one correspondence between the two clases of adverbs. "Style-disjuncts" (cf. Greenbaum (1969) and Schreiber (1972)), such as confidentially, honestly, and frankly in the following examples, occur in clause-initial and clause-final positions, but do not occur in pre-verbal positions without parenthetical intonation.

124. Confidentially, she's no friend of mine.
125. Honestly, I didn't mean to insult you.
126. Frankly, I simply don't like you.

Temporal adverbs, such as yesterday and tomorrow, also occur in clause-initial and clause-final positions, but not pre-verbal positions. Some frequency adverbs, such as occasionally and frequently, occur in clause-initial and clause-final positions, and also occur in pre-verbal positions. It can be assumed that these adverbs are in lexical class 2, but, as explained in section 4, it will still be necessary to posit a distinct category to which these adverbs belong when in pre-verbal positions. A few VP adverbs, such as quickly and slowly, also occur in clause-initial and clause-final positions. All VP adverbs occur in pre-verbal positions, but generally do not occur in pre-verbal position when the verb is an auxiliary. Thus, if quickly or slowly occur in pre-auxiliary position, then it is reasonable to assume that they occur in this position by virtue of belonging to lexical class 1. Speakers differ as to whether or not they accept these adverbs in pre-auxiliary positions; some speakers accept sentences such as 127 and 128, while others do not.

127. The man quickly will bang the drum.
128. The children slowly have recited the alphabet.

This difference can be accounted for if we assume that for some speakers, these adverbs are members of lexical class 1, but for other speakers, they are not.

FOOTNOTES

1. I am using the type assignments for NP's and VP's given in Klein and Sag (1982). In Klein and Sag (1982), VP's are third order predicates. Dowty (1979) and Bach (1980) have argued that modal and tensed VP's should be analyzed as third order predicates. As noted by Dowty (1980), once tensed and modal VP's are defined as third order predicates, "an expression of the category PredP [i.e. a tensed or modal VP] has the meaning of the subject of the sentence within its scope. Hence an adverb like possibly, which is part of the PredP, can likewise have the subject within its scope, which is the crucial semantic property that S-adverbs must have" (p.7).
2. Pullum (1982) gives arguments for the claim that to is a verb.
3. "Style disjuncts" (cf. Greenbaum (1969) and Schreiber (1972)), such as confidentially, honestly, and frankly in the following examples, have been labeled sentential adverbs, but do not occur in pre-verbal positions without parenthetical intonation.
   Confidentially, I wouldn't trust him.
   Honestly, I love that color on you.
   Frankly, I can't stand divinity fudge.
4. It should be noted that the rule of minor right node raising, [VP -> VP/NP NP], is not produced by the rightward displacement schema of Gazdar (1981), repeated below, because it is required that a be a clausal category.
   RDS: $\langle 9, [a \rightarrow a/B\ B],\ h_B[(a/B)'](B')\rangle$
   where a ranges over clausal categories
   and B can be any phrasal or clausal category.
5. Note that a sentence such as Ed said he would catch and kill the rabid dog and catch and kill the rabid dog, he will. cannot be used to motivate the minor right node raising rule, given the rule [V -> V and V] which Gazdar (1982) assumes. Given this rule catch and kill the rabid dog can be generated as a VP consisting of a complex V and an NP. Rather than argue against this complex V analysis, I have given an example in which it is clear that two verbs have not been conjoined, since try to kill is not a verb.
6. We cannot account for adverbs in clause-initial position by allowing the Topicalization Schema to apply to adverbs. If the Topicalization Schema were allowed to apply to adverbs, it would incorrectly be predicted that i and ii have readings in which the adverb has scope over only the lower clause.
   i. Probably, Jake knows that Howard will leave.
   ii. Unfortunately, Jake knows that his brother has been in an accident.
   Since the Topicalization Schema is stated as not applying to lexical categories, topicalization of adverbs is correctly ruled out. It is also necessary to rule out topicalization of adverb phrases. Otherwise, we predict an interpretation of iii in which the ADVP has scope over the lower clause.
   iii. Quite possibly, Jake knows that Howard will leave.
7. This example is from Jacobson (1964), citing B. Ifor Evans, A Short History of English Literature.

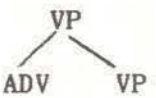
4. Temporal and Frequency Adverbs

In this section I will compare the placement of temporal and frequency adverbs with the placement of evaluative and modal adverbs. The results of this comparison will be important in the analysis of adverb stranding to be presented in section 6.

Adverbs which specify frequency, such as always, sometimes, occasionally and usually, occur in most of the same positions as evaluative and modal adverbs. Temporal adverbs, which specify a particular period of time, such as yesterday, today, and tomorrow, occur in much fewer positions.

Frequency adverbs, like evaluative and modal adverbs, occur in the configuration in 1 when immediately preceding a main or auxiliary verb.

1.



Examples such as 2 and 3 provide evidence for the higher VP node in 1.

2. Laurie said she would always love her mother, and always love her mother, she will. [VP Preposing]
3. Danny will always love Marsha, and Mark will, too. [VP Deletion]

If 2 and 3 are examples of VP Preposing and VP Deletion, as they certainly seem to be, then always love her mother and always love Marsha must be VP's. Examples such as 4 provide evidence for the lower VP node in 1. Because of the presence of sometimes in the second conjunct, the only reading of 4 is one in which only paid for dinner has been 'deleted' semantically, and not usually paid for dinner. If this is an example of VP Deletion, as it certainly seems to be, then paid for dinner in the left conjunct must be a VP.

4. Michael has usually paid for dinner, and Beth sometimes has.

Temporal adverbs such as yesterday and tomorrow do not occur in positions immediately preceding main or auxiliary verbs, as examples such as 5-7 show.

5. *John yesterday went to the beach.
6. *John will tomorrow go to the beach.
7. *John tomorrow will go to the beach.

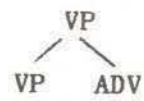
These adverbs, therefore, do not occur in the configuration in 1.

Temporal and frequency adverbs, unlike evaluative and modal adverbs, do occur at the right of conjoined VP's, as in 8 and 9.

8. Clark writes letters usually and sends telegrams sometimes.
9. Clark wrote a letter yesterday and sent a telegram today.

Such sentences can be accounted for if these adverbs occur in the configuration in 10.

10.



The following chart sums up the positions of occurrence of evaluative, modal, frequency, and temporal adverbs when not clause initial or clause-final. The phrase structure rules by which each type of adverb must be introduced are also listed.

	Left sister of VP	Right sister of VP	P-S rules
Evaluative and modal	yes	no	VP -> ADV VP
Frequency adverbs	yes	yes	VP -> ADV VP
			VP -> VP ADV
Temporal adverbs	no	yes	VP -> VP ADV

As shown, frequency adverbs occur as both left and right sisters of VP. Temporal adverbs occur only as right sisters of VP, while evaluative and modal adverbs occur only as left sisters of VP (cf. Section 3). In order to account for these observations, it is necessary to include the three ID rules and two LP rules given below.

11. <1, [VP -> VP, ADV] ADV'(VP')>
    where ADV = the evaluative and modals
    [1]
12. <2, [VP -> VP, ADV], ADV'(VP')>
    where ADV = frequency adverbs
    [2]
13. <3, [VP -> VP, ADV], ADV'(VP')>
    where ADV = temporal adverbs
    [3]
14. VP < ADV
    [3]
15. ADV < VP
    [1]

Given these ID and LP rules, the following phrase-structure rules will be basic rules of the grammar.

16. <1,[VP -> ADV VP], ADV'(VP')>
17. <2,[VP -> ADV VP], ADV'(VP')>
18. <2,[VP -> VP ADV], ADV'(VP')>
19. <3,[VP -> VP ADV], ADV'(VP')>

The ID rules in 11-13 cannot be collapsed in any way, because of the need to enforce different linear precedence restrictions on the three types of adverbs. I have used the rule numbers in the LP rules. Some other feature could have been used, but we would still need three distinct ID rules, since the adverbs in the ID rules would have to be marked with different features.

When Gazdar and Pullum (1981) introduced the ID/LP format, they did not specify any requirements on the form of LP rules. In Gazdar and Pullum (1982:21), it is claimed that the LP rule in 20 is a rule of English.

20. [+LEXICAL] < [-LEXICAL]

Thus, it is clear that Gazdar and Pullum intend features other than syntactic category features to be permitted in LP rules. Since "the rule number is assigned, by convention, to be the value of the feature LEXICAL" (Gazdar and Pullum (1982:17)), the use of rule numbers in LP rules is certainly not ruled out by the theory as stated. It should be noted that the rules in 18 and 19 are inconsistent with Gazdar and Pullum's generalization in 20. However, if we replace ADV with ADVP (adverb phrase) in these rules, they are no longer inconsistent with 20. I will leave open whether such a move should be taken, but note that if we replace ADV with ADVP in 16-19, we can account for the difference in positions of occurrence of these three kinds of adverbs by distinguishing three categories of ADVP, as well as three lexical classes of ADV's.

5. VP Adverbs

In this section an analysis of the placement of VP adverbs is presented. It is argued that VP adverbs occur as sisters to V, rather than, as Jackendoff (1972, 1977) claimed, as aunts of V.

I will limit this discussion to those VP adverbs with the fewest restrictions on their occurrence in positions other than clause-initial and clause-final. This group includes quickly, slowly, intensely, incessantly, thoroughly, seriously, firmly, diligently, completely, tremendously, purposefully, as well as willingly, knowingly, and cleverly.[1]

It is uncontroversial that VP adverbs are dominated by VP. It must be argued, however, that they occur as sisters, rather than aunts of V. Examples of 'VP Deletion' provide evidence for this claim.

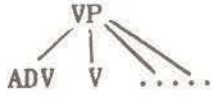
1. John has been seriously wounded, and Mary has been, too.
2. George has quickly read the book, and Mary has, too.
3. George had firmly refused the offer, and Mary had, too.

The only interpretation which these sentences have is one in which the adverb is included as part of what has been 'deleted' semantically. The only interpretation of 1, for example, is one in which seriously wounded has been 'deleted' semantically. In section 3.25 it was shown that a 'deleted' VP in a right conjunct could correspond to a VP in the left conjunct dominating an S-adverb and a VP or it could correspond to the VP which is a sister of the S-adverb. In 4, the most natural reading is one in which the VP dominating the ADV and following VP has been deleted. In 5, the only reading is one in which the VP which is a sister of the adverb has been deleted.

4. Two plus two will necessarily equal four and one plus one will, too.
5. John will probably go to Baltimore, and Mary definitely will.

If the adverbs in 1-3 were aunts of the V, we would expect these sentences to be ambiguous between an interpretation in which the higher VP has been 'deleted' semantically and one in which only the lower VP has been 'deleted'. Sentence 1 should be ambiguous between an interpretation in which seriously wounded is 'deleted' and one in which wounded is 'deleted'. However, this is not the case. This observation can be accounted for if there is no lower VP to which the adverb is a sister. In the following discussion, I will assume that VP adverbs are to be sisters of V, as in configuration 6.

6.



VP adverbs occur before main verbs, but not before auxiliary verbs, as the following examples show.

7. George quickly read the book.
8. *George quickly has read the book.
9. George has quickly read the book.

10. *George will quickly have read the book.
11. George will have quickly read the book.

Jackendoff (1972:75) marks examples such as 10 with a question mark, rather than an asterisk. However, my informants consistently rejected sentences such as 10. Some speakers accepted sentences such as 12 with quickly preceding the auxiliary verb, but only with an interpretation in which John was quick to bang the drums, not with the interpretation that the banging was quick.

12. John quickly has banged the drums.

VP adverbs also occur before prepositional phrases within the VP, as in 13, and in VP-final position, as in 14.

13. John gave the book quickly to Mary.
14. John gave the book to Mary quickly.

They do not occur before noun phrases within the VP, as 15-17 show, or before VP or S complements of the verb, as 18-20 show.

15. *John gave quickly the book to Mary.
16. *John gave quickly Mary the book.
17. *John persuaded quickly Kim to leave.
18. *John persuaded Kim quickly to leave.
19. *John wanted Kim quickly to leave.
20. *John promised Kim quickly that he would visit her.

In order to account for the data presented above, I will adopt the metarule in 21, as well as three LP rules.[2]

21. $\langle 8, [VP \rightarrow V, \underset{[-AUX]}{\ } X], V'(F)\rangle \Longrightarrow$
$\langle 8, [VP \rightarrow ADV, V, \underset{[-AUX]}{\ } X], ADV'(V'(F))\rangle$
where ADV = the VP adverbs

This metarule states that for every ID rule of the grammar which expands VP as V and any categories X, there is also an ID rule which expands VP in exactly the same manner except ADV is also a daughter of VP. This metarule captures the generalization that VP adverbs may occur as daughters of any VP. I have given a semantic translation in which the semantic value of the adverb is a function from verb phrase type denotations to verb phrase type denotatins. The ID rules which will be the output of the metarule in 21 will have two lexical categories, ADV and V, as daughters of VP. Rule numbers, by convention, are features on the lexical category introduced by an ID rule, but in these rules two lexical categories are introduced. In order to ensure that the ID rules which are output by 21 will have V, and not ADV, marked with the rule number, I will adopt a convention that only the lexical category which is marked with the rule number in the ID

rules which are output by 21 will have V, and not ADV, marked with the rule number. I will also assume that lexical categories which are not marked with a rule number by this convention may have a rule number specified as one of their features; otherwise it would not be possible to specify which adverbs occur in the ID rules output by metarule 21. Metarule 21 will be revised as in 22.

22. Revised VP metarule
    <8, [VP -> V, X], V'(F)> ==>
    <8, [VP -> ADV, V, X], (ADV'(V'))(F)>
    [6]
    where ADV = the set of VP adverbs
    [6]

In order to rule out ungrammatical sentences such as 18-20, it is necessary to adopt the LP rules in 23 and 24.

23. VP < ADV
    [6]
24. S < ADV
    [6]

In order to rule out ungrammatical sentences such as 15-17, but allow grammatical sentences such as 7, it will be necessary to adopt a new type of LP rule. Note that 15-17 cannot be ruled out by adopting the LP rule in 25, since this rule would incorrectly rule out sentences such as 7, in which the adverb precedes the NP.

25. NP < ADV
    [6]

What is needed is an LP rule which allows adverbs of class 6 to precede NP's, as in 7, but not to immediately precede NP's. LP rules, as originally conceived by Gazdar and Pullum (1981), cannot make a distinction between precedence and immediate precedence. By dis allowing such distinctions to be expressed by LP rules, the predictive power of the LP/ID format is enhanced. However, if the metarule analysis of VP adverbs is correct, it will be necessary to allow LP rules to make such a distinction. The rule in 26, which states that adverbs of class 6 may not immediately precede NP's will be needed.[3]

21. ADV I⁄ NP
    [6]

Footnotes

1. Willingly, knowingly, and cleverly, and other "passive-sensitive" adverbs, also belong to the same lexical class as evaluative and modal adverbs, since they occur in te same positions as these adverbs, as well as in the same positions as VP adverbs. Willingly, knowingly, and cleverly are known as "passive-sensitive" adverbs, because of the difference in interpretation of examples such as i and ii.
   i. The doctor willingly examined Mary.
   ii. Mary was willingly examined by the doctor.

The passive sentence ii has a reading which the active sentence does not: ii has a reading in which Mary was willing, as well as a reading in which the doctor was willing, but i only has the reading in which the doctor was willing. Given that these adverbs belong to the same lexical class as evaluatives and modals, as well as VP adverbs, it will be possible to adopt Dowty's (1980) analysis of the semantics of passive sensitive adverbs (cf. footnote 2 below).

2. We can follow Dowty's (1980) treatment of passive-sensitive adverbs if we incorporate transitive verb phrases (TVP's) into the grammar (cf. Gazdar and Sag (1981) for a discussion of TVP's in GPSG) and adopt the metarule below in addition to metarule 21. (These two metarules could be collapsed into a single metarule.)

<10, [TVP -> V, X], V'(F)> ==>
<10, [TVP -> ADV, V, X], ADV'(V'(F))>
[9]

In Dowty's analysis passive-sensitive adverbs belong to the class PredP/PredP (which corresponds semantically to our lexical class 2), IV/IV (which corresponds to our lexical class 6), and TV/TV (which corresponds to our lexical class 9).

3. It has been pointed out to me by David Dowty that this is not the only possibility. The same data could also be accounted for by permitting rules consisting of disjunctions of LP rules. The rule below would yield the same results as rule 26.

ADV < V or ADV > NP VP, S
[6] [6]

6. Adverb Stranding

This section deals with sentences in which an adverb immediately precedes a 'deletion' or 'movement' site, and is thus stranded. It will be shown that given the rules for adverb placement proposed in previous chapters, and the Trace Introduction Metarule of Sag (1982), data involving adverb stranding is immediately accounted for.

Quantifiers, like sentential adverbs, cannot immediately precede 'deletion' or 'movement' sites. The analysis to be presented, unlike previous analyses, will account for the identical 'behavior' of sentential adverbs and quantifiers before movement and deletion sites.

In section 6.1 previous analyses of adverb (and quantifier) stranding will be discussed. In section 6.2 a rule will be proposed for quantifier placement and the Trace Introduction Metarule will be discussed. In section 6.3, it will be shown that the Trace Introduction Metarule interacts with the rules for adverb and quantifer placement to yield correct predictions concerning sentences in which adverbs and quantifiers immediately precede VP 'movement' sites. In this section, it will also be shown that, given certain assumptions about the feature null, correct predictions result concerning adverbs and quantifiers before VP 'deletion' sites.

6.1 Previous analyses

Baker (1981) discusses the ungrammaticality of sentences such as 1 and 2 in which a sentential adverb or quantifier precedes a deletion site.

1. *Fred has never been rude to Grandfather,
   but John has always ∅.
2. *I have read Moby Dick, and they have all ∅, too.

Baker assumes that adverbs and quantifiers precede the finite auxiliary verb in underlying structure. A transformational rule of Auxiliary Shift moves unstressed auxiliaries to the left of adverbs or quantifiers. The sentence in 3, for example, derives from 4 by Auxiliary Shift. Auxiliary Shift moves the stressless auxiliary have to the left of the adverbs probably and never.

3. George and Martha have probably never seen a real politician.
4. George and Martha probably never have seen a real politician.

According to Baker's analysis, the underlying structures for 1 and 2 would have to be 5 and 6, respectively.

5. Fred has never been rude to Grandfather,
   but John always has ∅.
6. I have read Moby Dick, and they all have ∅, too.

Baker claims that auxiliaries before deletion sites are always stressed. Since the auxiliary is stressed, Auxiliary Shift will not apply to 5 or 6, and thus 1 and 2 are predicted to be ungrammatical.

Baker's analysis hinges on the claim that only unstressed auxiliaries occur before adverbs or quantifiers. Sag (1980) and Ernst (1983) have cited the following counterexamples to this claim.

7. They denied that John has always admired Susan, but he HAS always admired her. (Sag's example 6)
8. a. Do you mean to say in front of this committee, sir, that every single factor has been taken into account in your budget estimates?
   b. Well...we HAVE probably glossed over the effects of the FOOD PRICE increases. (Ernst 1983)

In 7 the auxiliary preceding the adverb in the second clause is stressed and the VP of the first clause is 'echoed' to some extent. In 8.b. the adverb follows a stressed auxiliary and a second stress occurs in the VP. Ernst (1983) discusses various discourse conditions under which sentences such as 8.b. are acceptable. He suggests that "adverbs may follow auxiliaries whenever discourse conditions allow the auxiliaries to be stressed--whether or not there is a second stress in the VP" (p.547). Thus, according to Ernst, the acceptability of sentences such as 7, as well as sentences such as 8.b., is dependent on the discourse situation. He concludes that "we should allow the grammar to generate adverbs freely after stressed auxiliaries, in addition to the regular cases of nonstressed auxiliaries. The only requirements are discourse conditions relating to appropriate structures for contrasts and to different degrees of stress" (p.547).

It is clear that Baker's analysis is incorrect and that the relation between auxiliary stress and adverb placement is governed by discourse conditions. It is not clear how to rule out sentences such as 1 and 2, which are ungrammatical whether or not the auxiliary is stressed.

Sag (1978, 1980) considers the ungrammaticality of such sentences to be related to the ungrammaticality of the sentences below. According to Sag, these sentences are ruled out by the generalization that "adverbs and so-called 'floated' quantifiers may not appear in surface structure in a position immediately preceding an extraction site" (1980:255).

9. *I don't know what they are all ø.
   *I don't know how happy they are ever ø.
   [WH movement]
10. *I know a first grader who has finished more lesson units than the second graders have all ø.
   *The activists are now more active than they were ever ø.
   [Comparative Deletion]
11. *My brother has studied karate, and my sisters have all ø, also.
   *I don't know if Leslie has ever studied karate, and I don't know if Gwendolyn has ever ø, either.
   [VP Deletion]
12. *Sandy is polite to strangers, which I doubt very much that your brothers are all ø.
   *Sandy is polite to strangers, which I doubt very much that Ralph is ever ø.
   [Relativization]

13. *None of them were Communists, but Socialists, they were all ∅.
*They used to be Socialists, but Communists, they were never ∅.
[Topicalization]

14. *The more unhappy you say they are, the happier they are all ∅.
*The more polite you tell them to be, the more polite they are usually ∅.
[The-More-the-Merrier-Fronting]

15. *They said our children would be polite and polite they are all ∅.
*They said our children would be polite, but polite, they are never ∅.
[VP Preposing]

The surface filter in 16 is posited to account for the ungrammaticality of 9-15.

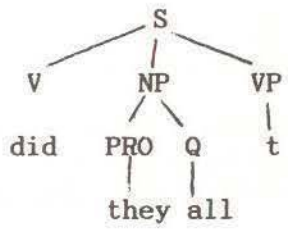
16. *{Q, ADV} extraction site

Sag (1978) notes that this surface filter must be modified in view of the grammaticality of questions such as 17 and 18.

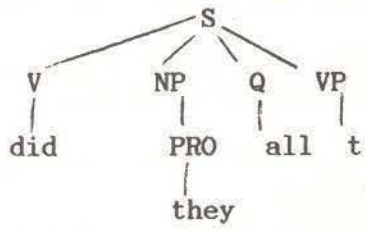
17. Did they all ∅?
18. Does he usually ∅?

He sketches a solution to the problem presented by questions such as 17. Such questions, he claims (following Postal (1974) and Maling (1976)) have two constituent structures: one in which PRO and Q form a constituent (NP), as in 19, and another in which they do not, as in 20.

19.



20.



In 19 Q is not the sister of an extraction site, but in 20 it is. Sag accounts for the grammaticality of questions such as 17 by revising the filter as in 21.

21. * { Q / ADV } extraction site

where Q, ADV is a sister of the extraction site

This filter rules out an adverb or quantifier before an extraction site only if it is a sister of the extraction site. Sag obviously intends "sister of an extraction site" to be taken to mean sister to a node which immediately dominates a trace or null element. This filter will not rule out the grammatical sentence in 17, because 17 has a structure, 19, in which the Q is not a sister to the extraction site. Note that sentences such as 22 are correctly ruled out by the filter in 21 because they have only a constituent structure in which the Q is a sister of the extraction site.

22. *Did the men all?

Sentences such as 18 will still be incorrectly ruled out, however, since there is no evidence for a constituent structure in which the adverb is not a sister of the extraction site. Sag (1978) claimed that examples such as 18 were grammatical only if a pronoun precedes the adverb, and cited the following examples which contrast with 18.

23. *Does President Carter usually ∅?
24. *Will Anita Bryant ever ∅?

However, my informants judged examples such as 25 to be perfectly acceptable. Perhaps the unacceptability of 23 and 24 has to do with the use of proper names.

25. Do your friends usually?
(with no pause before usually)

Examples such as 26, from Baker (1981), will also be incorrectly ruled out by Sag's filter, as well as 27 and 28 from Ernst (1983).

26. He's gotten along well with Fred in the past few weeks,
but he hasn't always.
27. Terry knows how to build an H-bomb.
No--does he REALLY??
28. Joe says he will run a four-minute mile on a steeple-
chase course.
How could he POSSIBLY?!!

Ernst (1983) notes that the counterexamples to Sag's surface filter involve a restricted set of adverbs: time adverbs, such as usually, sometimes, then, now, recently, soon, and the two adverbs really and possibly (for some speakers). However, there is another type of counterexample in which VP adverbs, such as quietly, partially, and slowly, apparently are sisters to deletion sites. These examples involve the verbal ellipsis phenomenon known as "gapping".

29. John will loudly answer my questions and
    Mary ø quietly ø.
30. Todd has thoroughly read the book and
    Mark ø partially ø.

In examples 29 and 30, the auxiliary, the verb, and object NP are 'missing'. Whether the VP adverbs are sisters to a VP which dominates the verb and object or are within the VP and sisters to the V, they are sisters of extraction sites. If the adverb is a sister to VP, one might consider arguing that 29 and 30 are actually examples of VP Deletion--that, for example, will answer my questions has been 'deleted' before quietly and nothing has been 'deleted' after the adverb. However, evidence can be given to the contrary. VP Deletion can apply within an embedded S, as in 31, but gapping cannot; although 32 is grammatical, the sentences in 33 (from Sag 1977) are not.

31. John will go to the movies and I know that Bill will, too.
32. Alan went to New York and Betsy ø to Boston.
33. *Alan went to New York, and
    a. I know (that)
    b. it seems (that)
    c. Bill met a man who claimed (that)
    Betsy ø to Boston.

If 29 and 30 were examples of VP Deletion, we would expect sentences such as 34 and 35 to also be grammatical, but they are not.

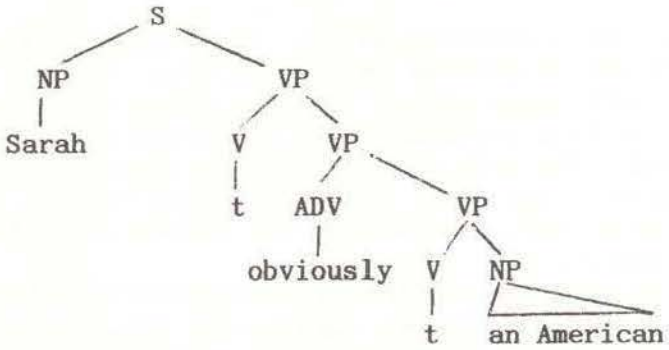
34. *John will loudly answer my questions and I know that
    Mary quietly.
35. *Todd has thoroughly read the book and I know that
    Mark partially.

The grammaticality of sentences such as 29 and 30 can be accounted for if it is assumed that the surface filter applies only to sentential adverbs. (Ernst seems to assume that this is what Sag intended anyway.) There are similar examples of Gapping involving sentential adverbs, but in these cases the adverbs are not sisters of the extraction sites.

36. Olga will probably marry a Russian and Sarah ø obviously
    ø an American.

Given the analysis of modal adverbs in chapter 3, the right conjunct in 36 will be assigned the structure in 37.

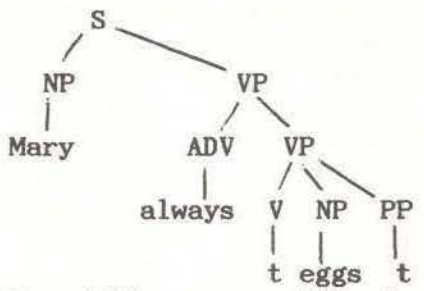
37.



The sentence in 36 is a counterexample to the original filter in 16. It is not a counterexample to the revised filter in 21, however, since obviously is not a sister of the extraction site.1 There are also examples of frequency adverbs preceding 'gapped' verbs, as in 38. I have argued that such adverbs are sisters of VP. If this is the case, then the adverb in 38 is not the sister of an extraction site, but but instead the aunt, as in 39.

38. John usually eats cereal for breakfast and
Mary always ø eggs ø.

39.



Given that 36 and 38 are examples of gapping, they count as evidence against the original filter in 16, but not the revised filter in 21.

To sum up the discussion of filters: the following counter-examples to the original filter have been given:

40. Do they all?
41. Olga will probably marry a Russian and Sarah ø obviously ø an American.
42. John usually eats cereal for breakfast and Mary always ø eggs ø.
43. ...does he REALLY?
44. ...How could he POSSIBLY?!!
45. Does he usually?
46. He's gotten along well with Fred in the past few weeks, but he hasn't always.
47. John will loudly answer my questions and Mary ø quietly ø.
48. Todd has thoroughly read the book and Mark ø partially ø.

Once the surface filter is revised, as in 21, the examples in 40-42 will no longer be counterexamples. However, the examples in 43-48 are apparent counterexamples to the revised filter. I will have nothing to say about 43 and 44. The other examples involve either time adverbs or

VP adverbs occuring before deletion sites. The correct observation seems to be that sentential adverbs (excluding time adverbs) cannot occur as sisters of extraction sites. While a surface filter can be devised to account for this observation, I would claim that such an analysis is misguided. Not only does it fail to explain why Q's and S-ADV's should 'behave' alike (i.e. why both should fail to occur as sisters of extraction sites), but, within the framework to be presented, it is unnecessary.

Baker (1981) and Ernst (1983) have both claimed that Sag's surface filter analysis "seems rather implausible from the point of view of language acquisition" (Ernst, p. 547). Ernst proposes that Sag's filter be replaced by a filter which forbids material between auxiliary verbs and a VP-deletion site. However, as Ernst points out,such a filter incorrectly rules out examples of Subject-Auxiliary Inversion such as that in 49.

49. Phil was diving into a wet dishrag.
WAS he ø?! (Ernst 1983)

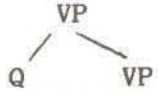
In the following sections, an analysis will be offered within a version of Generalized Phrase Structure Grammar. In this analysis, no special constraint is needed to rule out ungrammatical sentences in which a sentential adverb (excluding time adverbs) or a quantifier is a sister of an extraction site (actually what is to be excluded is a structure in which the node S-ADV or the node Q is a sister of an extraction site). Certain independently motivated aspects of the grammar interact to produce the desired results. Since no surface filter will be required, this analysis is compatible with Jacobson's (1982) tentative claim that "no constraint in the grammar can explicitly mention gaps" (p. 207). Under a filter analysis it just happens to be the case that both Q's and S-ADV's cannot be sisters of extraction sites. Under the analysis to be presented, Q's and S-ADV's 'behave' alike in this respect because of a structural identity.

6.2 Assumptions Underlying the Structural Analysis

In the structural analysis to be presented, I will assume the rules for introducing S-ADV, frequency, temporal, and VP adverbs given in the preceding chapters.

I will also assume that 'floated' quantifiers occur in the configuration in 50. Baltin (1982) argues for such a structure.

50.



Examples of VP Preposing and VP Deletion provide evidence for the higher VP node in 50.

51. They said that they would all work on that, and all work on that, they did. (Baltin 1982, example 36)
52. They said they will all work on that and they will.

Examples such as 53 provide evidence for the lower VP node in 50.

53. The women will all go to Rapid City and Howard will, too.

Quantifiers in pre-verbal positions will be introduced by the basic rule in 54.

54. <10, [VP -> Q VP]>
where Q(10) = all, each, both

I have not given a semantic rule in 75. It is necessary to provide a semantic analysis of sentences with 'floated' quantifiers which is compatible with the syntactic rule in 54, if this analysis is to be viable. However, since the evidence for the syntactic rule in 54 is compelling, I will assume for now that such a corresponding semantic rule can be motivated.

In the analysis to be presented it will also be assumed that all traces are introduced by means of the Trace Introduction Metarule (TIM) of Sag (1982). The TIM does much the same work for GPSG which the Immediate Dominance Principle of Sag (1977) did for TG. The TIM replaces linking rules such as 55.

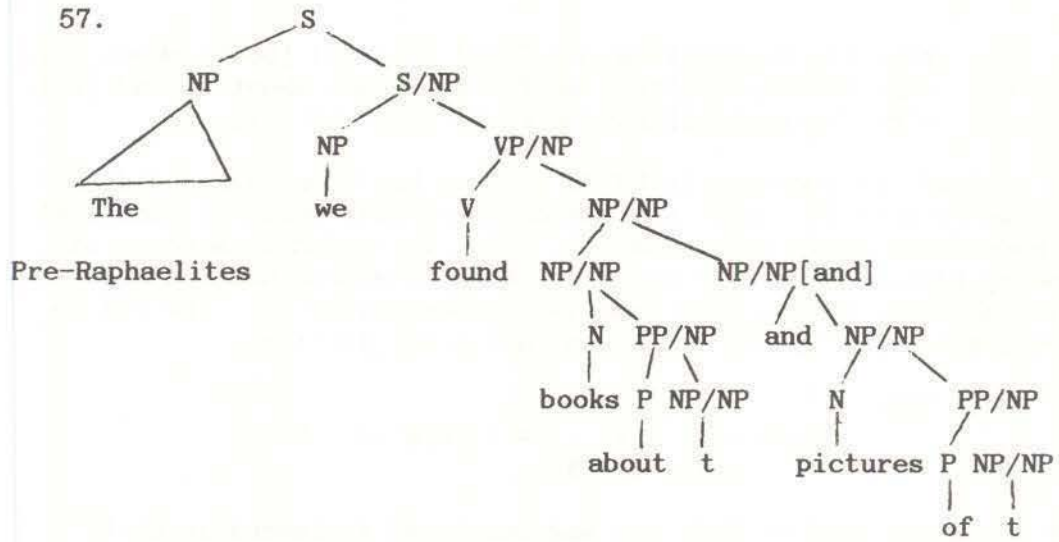
55. [NP/NP -> t]

The TIM was proposed in order to avoid problems for Gazdar's (1981) treatment of coordination which were due to the use of linking rules such as 55. Gazdar's coordination schema in 56 allows for coordination of NP/NP's as in 57.

56. <2, [a -> $a_1 \ldots a_n$], B'(a_1',...,a_n')>
[B]
where B ∈ [and, or] and a is any syntactic category

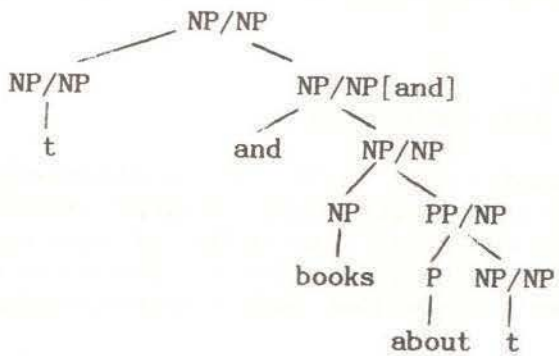
<3, [a -> B a], a'>
where B ∈ [and, or,...] and a is any syntactic category

57.

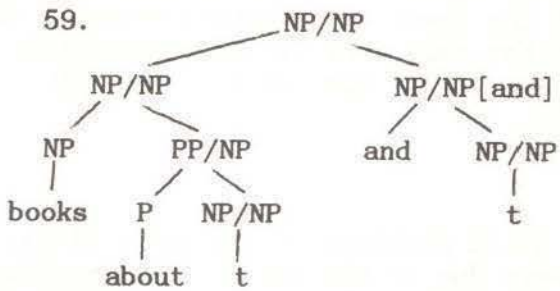


Given the linking rule in 55 and Gazdar's coordination schema, subtrees as in 58, 59, and 60 (Sag's 13 a,b,c) will be allowed and the ungrammatical sentences in 61, 62, and 63 (Sag's 14 a,b,c) will be generated.

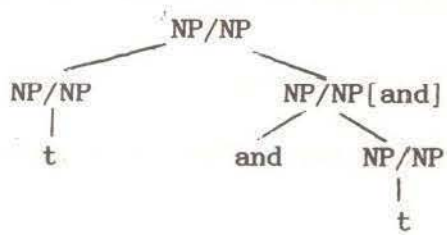
58.



59.



60.



61. *The Pre-Raphaelites, we found [[t] and [books about t]].
62. *The Pre-Raphaelites, we found [[books about t] and [t]].
63. *The Pre-Raphaelites, we found [[t] and [t]].

In each of the subtrees in 58-60 a trace has been introduced (by the linking rule in 55) under a slash category node which is identical to the node immediately dominating it. What is needed is a means of introducing traces which will not allow them to appear under a node which is identical to the node immediately dominating it. The TIM in 64 accomplishes this by the condition that a not equal B.

64. TIM:

$$[a/B \rightarrow \ldots B/B \ldots] \Longrightarrow [a/B \rightarrow \ldots t \ldots]$$
$$\text{where } a \neq B$$

If 55 is replaced by TIM, the ungrammatical sentences in 61-63 will no longer be generated. Sag (1982:333) states that 61 will not be generated because "TIM would have to produce rules like the one in (17) [65 below]."

65. [NP/NP -> t NP/NP]
[and]

"This could only happen if the coordination schema ... were taken as input to TIM. But on independent grounds (see Gazdar (in press)), metarules may not operate on nonfinite schemata" (p.313). But note that even if the TIM were allowed to operate on the coordination schema, rule 65 would not be produced by TIM. Rule 65 could only result if the TIM took 66 as its input, but 66 is not a possible input to TIM since the condition that a does not equal B is not met.

66. [NP/NP -> NP/NP NP/NP]
[and]

Sentences 62 and 63 are ruled out by the a ≠ B condition. As Sag states, to generate 62, "TIM would have to apply so as to produce the rule in (18) [67 below]" (p.333).

67. [NP/NP -> and t]
[and]

"However this is impossible, as the input rule here would be the rule in (19) [68 below] ... which violates the a ≠ B condition on TIM" (p.334)

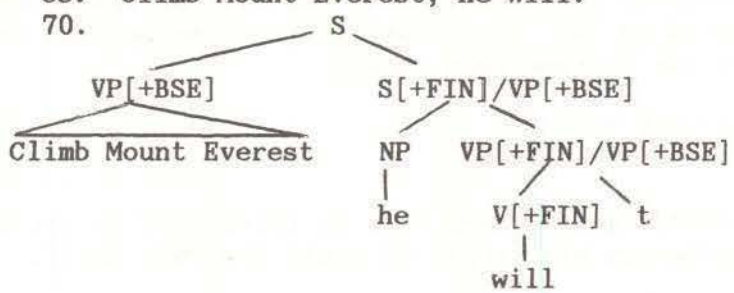
68. [NP/NP -> and NP/NP]
[and]

To generate 63, "one would need both rules (17) [67 above] and (18) [68 above] (p.334), which are, of course, ruled out."

Sag does not specify exactly what is meant by the condition a ≠ B. In the cases Sag discusses a ≠ B could be taken simply to mean that a is not the same category as B. However, such a condition will lead to incorrect predictions concerning VP Fronting. Given the analysis of VP Fronting presented in Gazdar et al. (1982), the tree for the sentence in 69 will be 70.

69. Climb Mount Everest, he will.

70.



The node dominating the trace in all cases of VP Fronting would be VP/VP. If the a ≠ B condition is taken to mean only that a and B may not have the same syntactic category features, then VP Fronting would be ruled out.

The TIM can be reconciled with VP Fronting, if we assume that a equals B only if the two nodes a and B are identical with respect to all features, the major category class simply being one of these features. Thus, a does not equal B if one or more features differ. Given this interpretation of TIM, VP Fronting will be allowed since the two VP's of the VP/VP node will differ in their features.

6.3 VP 'Deletion' and 'Movement'

In this section I will deal with cases in which an adverb or quantifier precedes a VP extraction site. These cases include examples of Comparative Deletion, VP Topicalization (VP Fronting), The-More-the Merrier Fronting, and VP Deletion. I am assuming that adjective phrases following the copular be are VP's marked as [+PRED], as do Gazdar et al. (1982). I will restrict the discussion to VP Preposing and VP Deletion, since it is not clear how examples of Comparative Deletion or The-More-the-Merrier Fronting should be handled.

6.31 VP Fronting

In Gazdar et al. (1982) VP Fronting is accounted for by the Topicalization schema below in conjunction with the slashing mechanism presented in Gazdar (1981) to account for unbounded dependencies.

71. <13, [S -> a S/a], λh_a [(S/a)'] (a')
when a = VP, then a is to be [-FIN, -INF, -ASP]

What is relevant to our discussion is the syntactic rule in the schema in 71 (i.e. [S -> a S/a]). The syntactic rule states that an S may consist of a phrasal category a followed by an S which is 'missing' an a. The slashing mechanism ensures that the VP which is 'missing' from S has the same features as the VP which is topicalized. Given the assumption that adverbs and quantifiers are 'Chomsky-adjoined' to VP's and that traces are introduced by TIM, I will show that the ungrammaticality of sentences such as 72 and 73 is predicted.

72. *John said he would pay me and pay me he will definitely ø.
(with no pause between will and definitely)
73. *They said they would all pay me and pay me they will all ø.

As in chapter 3, S-Adverbs in pre-verbal positions will be introduced by the basic rule repeated in 74. The slashing mechanism applies to this basic rule to give the derived rule in 75.

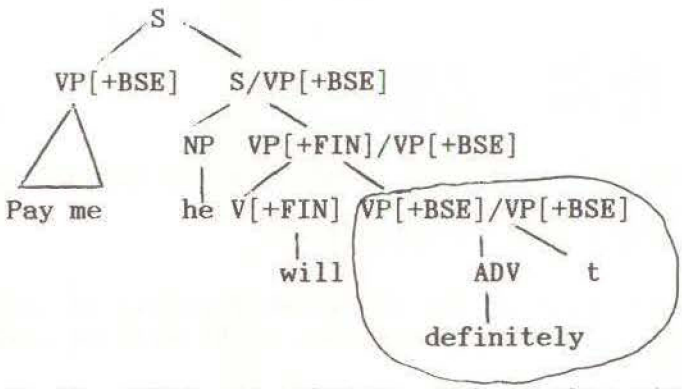
74. <1, [VP -> ADV VP]>
75. <1, [VP/VP -> ADV VP/VP]>

As stated earlier, 'floated' quantifiers will be introduced by the rule in 76. The slashing mechanism will apply to yield the rule in 77.

76. <10, [VP -> Q VP]>
77. <10, [VP/VP -> Q VP/VP]>

The tree for 72 would have to be 78. (The distribution of features is explained below.) The sentence in 73 would have the same structure.

78.



But the TIM, given in 64, rules out subtrees such as that circled in 78, since traces are only allowed as daughters of nodes a/B where a does not equal B. In the circled subtree the node immediately dominating the trace is an a/B where a is identical to B (i.e. a = B). Given the TIM, the tree in 78 is not a possible structure.

It can be shown that in any instantiation of the rule [VP/VP -> ADV VP/VP] in which the VP's of the dominated VP/VP node are identical, the VP's will all share the same features, and thus a will always equal B. To see why this is so, consider the subscripted version of the derived rule below.

79. $[VP_1/VP_2 \rightarrow ADV \quad VP_3/VP_4]$

The Head Feature Convention will ensure that VP_1 and VP_3 have identical features, since VP_3 is the head of VP_1. The TIM will require that VP_3 and VP_4 have the same features if a trace is to be introduced at this node (they would be the B/B in the TIM). Since VP_1 has the same features as VP_3 which has the same features as VP_4, VP_1 must have the same features as VP_4. The slashing mechanism will require that VP_2 and VP_4 have identical features. Thus VP_2 and VP_1 both have the same features as VP_4 and are therefore identical.

Given the TIM, then, it is impossible for a trace to appear as the sister of an ADV or Q which is introduced by one of the basic rules in 74 or 75 or one of the derived rules in 75 or 76. It is, therefore, predicted that S-Adverbs and quantifiers will not immediately precede the site of a 'moved' VP.

In section 5 I argued that VP adverbs are introduced by the metarule repeated in 80. Given this assumption, it is obvious why VP adverbs cannot immediately precede a 'moved' VP--they do not have VP's as sisters.

80. <n, [VP -> V, X], V'(F)> ==>
<n, [VP -> ADV, V, X], ADV'[V'(F')]>

Frequency adverbs apparently occur before 'moved' VP's as in 81.

81. John said he would always love his mother and love his mother, he will always.
(without a pause between <u>will</u> and <u>always</u>)

In section 4 I argued that frequency adverbs are introduced by the basic rules in 82 and 83.

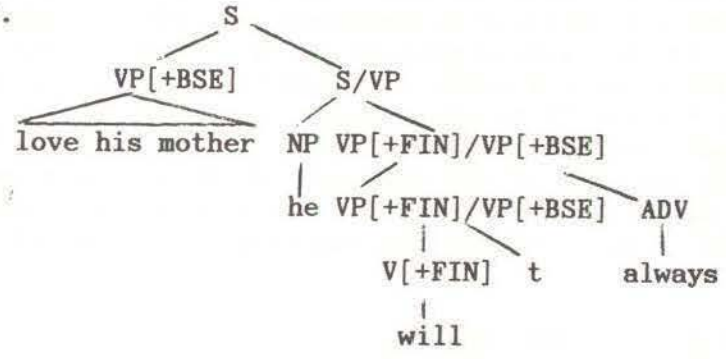
82. <3, [VP -> VP ADV]>
83. <3, [VP -> ADV VP]>

The slashing mechanism applies to 82 to yield the derived rule in 84.

84. <3, [VP/VP -> VP/VP ADV]>

Given the derived rule in 84, the grammaticality of sentences such as 81 is predicted. The second conjunct in 81 will be assigned the structure in 85.

85.



The slashing mechanism applies to the basic rule in 86 (cf. Gazdar et al. (1982)) to yield the derived rule in 87.

86. [VP -> V VP]
[+FIN] [+BSE]
87. [VP / VP -> V VP / VP]
[+FIN][+BSE] [+BSE][+BSE]

The rule in 87 can serve as input to the TIM since a does not equal B (i.e. VP[+FIN] is not identical to VP[+BSE]) and the rule in 88 will result.

88. [VP/VP -> V t]
[+FIN][+BSE]

Thus, the rules in 84 and 88, the slashing mechanism, and the Topicalization schema will interact to predict the grammaticality of sentences, such as 81, in which a frequency adverb immediately precedes a VP 'movement' site.

Since temporal adverbs are also introduced as right sisters of VP, we would expect sentences such as 89 to be grammatical, but they are not.

89. *John said he would go to the store tomorrow, and go to the store, he will tomorrow.
(with no pause before the adverb)

It is not clear to me how to explain the ungrammaticality of such sentences. If we are to maintain that temporal adverbs are introduced by the syntactic rule [VP -> VP ADV], as argued in chapter 4, it will

be necessary to provide an account of this data. If temporal adverbs were instead introduced by the rule [S -> S ADV], as evaluative and modal adverbs are, the ungrammaticality of 89 would be accounted for. However, it then becomes difficult to explain why these adverbs are not preceded by a pause as evaluative and modal adverbs are.

To summarize this section, the basic rules and TIM, along with other motivated rules of the grammar, interact to give correct predictions concerning quantifiers, S-Adverbs, frequency adverbs, and VP adverbs before VP 'movement' sites. It is not yet clear how to account for predictions involving temporal adverbs before VP 'movement' sites.

6.32 VP Deletion

S-Adverbs and VP adverbs do not occur before 'deleted' VP's, as the following examples illustrate.

90. *John has probably gone to Cleveland and
Mark has probably ø, too.
(with no pause between the auxiliary and adverb)
91. *Karen has thoroughly read this book and Doris
has thoroughly ø, too.

Quantifiers do not occur before 'deleted' VP's, unless immediately preceded by a pronoun.

92. *The men have all left for lunch and the women
have all ø, too.
93. *Have the men all ø?
94. Have they all ø?

Frequency adverbs apparently appear before 'deleted' VP's, whether or not a pronoun precedes. (I will argue below that the adverb is actually following the deleted VP, not preceding it.)

95. John has been nice to me lately, but he hasn't always.
96. Does he usually?
97. Do your friends usually?

As claimed in chapter 4, temporal adverbs do not precede VP's, and thus do not precede VP 'deletion' sites.

Gazdar et al. (1982) give the following metarules to account for 'VP Deletion':

98. VPD: <[VP -> V VP], F > ==>
[+AUX]
[-PRP]
[-GER]

<[VP -> V VP], F >
[+NUL]

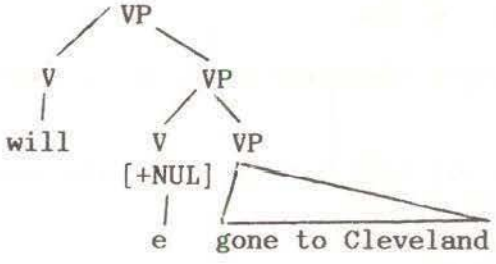
The metarule in 98 "takes any V [+AUX, -PRP, -GER] rule which expands as V followed by V', and simply adds the feature +NULL to the complement V'" (Gazdar et al. p. 606). The rule in 99 introduces e, which represents the empty string.

99. <16, [VP -> e], v>
[+NUL]
where v is a contextually bound variable ranging over VP denotations.

It is important to determine what kind of feature null is. Gazdar and Pullum (1982) discuss feature instantiation principles and distinguish two types of features--head features and foot features. Null is obviously not a head feature. If it were, the V of a [+null] VP would also be [+null] (since V is the head of VP and, by the Head Feature Convention, heads must have head features identical to those of their mother node), and ungrammatical sentences such as 100 would be produced. The right conjunct in 100 will be assigned the structure in 101.

100. John will have gone to Baltimore and
Betty will e gone to Cleveland.

101.



Since null is not a head feature, it may be a foot feature. Gazdar and Pullum (1982:34) note that "there are foot features that are explicitly specified in listed ID rules, or which have arisen through the operation of metarules. Such foot features are inviolate and cannot be copied or otherwise tampered with in the feature instantia tion mapping." What is important for our purposes is the claim that foot features which are explicitly specified in the syntactic rules are not subject to the Foot Feature Principle, given below.

102. Foot Feature Principle
The increment of the mother category's FOOT feature is the unification of the increments of the daughter categories' FOOT features.
(Gazdar and Pullum (1982:35))

This principle ensures, among other things, that all foot features of daughter nodes will also be features of the mother node.

The feature null is explicitly specified in the syntactic rule on the righthand side of the arrow in VPD metarule (98). Thus, it appears to be one of the foot features which does not obey the Foot Feature Principle. Implicit in Gazdar and Pullum's claim that foot features specified in syntactic rules "cannot be copied or otherwise tampered with in the feature instantiation mapping" is the assumption that the default value for features such as null is minus. All VP's will be [-null] by default, except the VP's specified as [+null] (i.e. VP's introduced by the VPD metarule). Thus, the VP's in the basic

rules for S-Adverbs and quantifiers will be [-null] by default. Since the VP immediately following the S-Adverb or Q is [-null], rule 99 cannot rewrite this VP as the empty string e. The ungrammaticality of sentences such as 90, 92, and 93 is thus accounted for.

The contrast between 93 and 94 is readily accounted for if we assume, as do Postal (1974), Maling (1976), and Sag (1978), that sentences such as 94 have at least one structure in which the Q and Pro form an NP and that sentences such as 93 only have a structure in which Q is part of the VP. This difference in structure is supported by the parenthetical test for constituent structure. Postal (1974) and Maling (1976) both note the following contrasts.

103. Malcolm proved them all, {don't forget / he claimed}, to be vicious criminals.

104. *Malcolm proved the soldiers all, {don't forget / he claimed}, to be vicious criminals.

I will assume the basic rule in 105, as well as the rule which 'Chomsky-adjoins' Q's to VP's.

105. [NP -> PRO Q]

The sentence in 94 involves Subject-Auxiliary Inversion which Gazdar et al. (1982) handle with the following metarule:

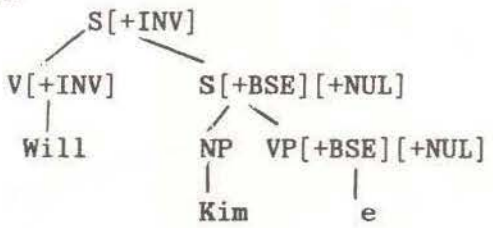
106. SAI:

[VP -> V	VP],	λ P[V'(V'(P))] ==>
[+FIN]	[a]	
[+AUX]		
[S -> V	S],	V'(S')
[+INV]	[a]	

They claim that "the VPD metarule ... feeds the SAI metarule" (p. 611) and that the sentence in 107 will be assigned the structure in 108.

107. Will Kim?

108.



The output of the VPD metarule serves as input to SAI to allow the rule on the righthand side of the arrow in 109.

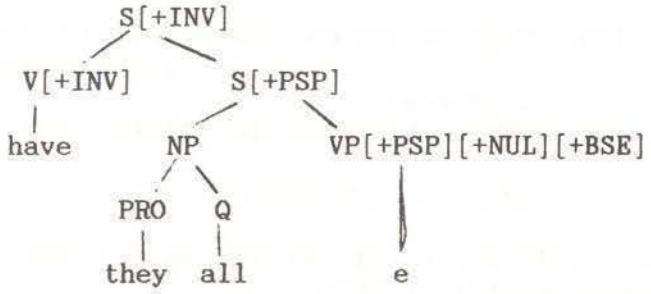
109. [VP -> V VP] ==> [S -> V S]

The tree in 108 makes it clear that Gazdar et al. are assuming that null is a head feature and that [+null] appears as a feature of V in 108 by virtue of the Head Feature Convention. But, as stated earlier,

assuming that null is a head feature will lead to the generation of ungrammatical sentences.

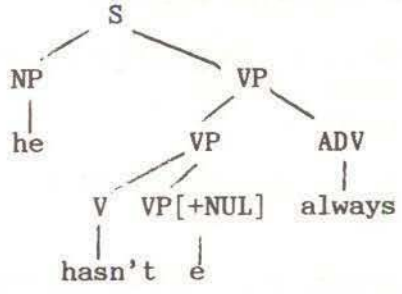
Obviously some other way of accounting for such sentences is needed. It is necessary to somehow specify that the VP in such structures may be [+null] without allowing for the generation of ungrammatical sentences. It is not obvious how this should be done. What is important for our purposes is that whatever means is used to account for such questions will also account for questions such as 93, assuming that the Q and PRO are both dominated by NP as in 110.

110.

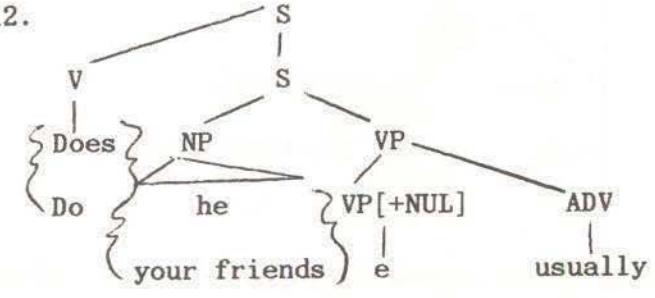


Because frequency adverbs can be 'Chomsky-adjoined' to the right of VP's, the grammaticality of the examples in 95-97 is correctly predicted. They will be assigned the following structures.

111.



112.



Sentences such as 113 below, with temporal adverbs will be given a similar structure.

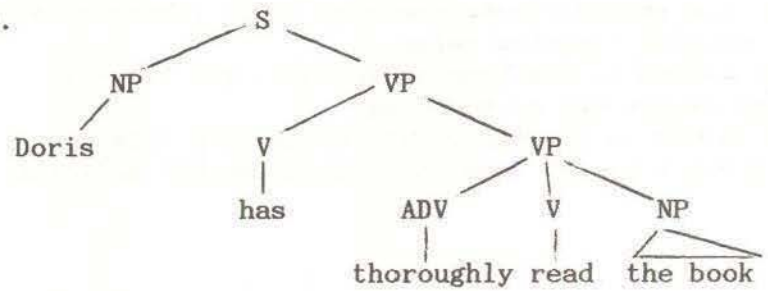
113. John will go to class today, but he won't tomorrow.

Since these adverbs are also introduced as right sisters of VP (cf. chapter 4), the grammaticality of such sentences is predicted.

The ungrammaticality of sentences such as 91 in which a VP adverb apparently precedes a VP deletion site is accounted for, because VP adverbs never immediately precede a VP. The structure for a sentence such as 114 will be 115. If the lower VP were 'deleted', the VP adverb would also have to be 'deleted'.

114. Doris has thoroughly read the book.

115.



It has been shown in this section that the facts about quantifier and adverb 'stranding' before VP deletion sites are readily accounted for given our assumptions about constituent structure and the treatment of VPD in Gazdar et al. (1982).

4.0 Advantages of the structural analysis

The structual analysis which has been presented to account for adverb and quantifier stranding facts is preferable to previous analyses for several reasons:

i. The identical 'behavior' of sentential adverbs and quantifiers before VP movement and deletion sites is explained. Quantifiers and sentential adverbs cannot precede VP extraction sites because they are 'Chomsky-adjoined' to VP's.

ii. The contrast between sentences such as 100 and 101 follows from the difference in structures these questions may have.

116. *Did the men all?

117. Did they all?

iii. The grammaticality of sentences in which frequency adverbs immediately precede VP extraction sites is accounted for.

iv. I have not discussed quantifiers and adverbs before NP extraction sites. However, assuming that quantifiers (Q) and sentential adverbs (S-ADV) are 'Chomsky-adjoined' to NP's, the ungrammaticality of sentences in which a quantifier or sentential adverb immediately precedes an NP extraction site is accounted for in the same way as the ungrammaticality of sentences in which a quantifier or sentential adverb immediately precedes a VP extraction site.

FOOTNOTES
SECTION 6

1. It might be claimed that 36 cannot be an example of gapping, since it is sometimes assumed that only two remnants may be left behind by gapping. However, there are other examples in which three constituents remain. Sag (1977:144) points out cases where the gapped clause contains three remnants (NP-PP-PP), as in his examples repeated below.
   i. Peter talked to his boss on Tuesday, and Betsy ø to her supervisor on Wednesday.
   ii. John talked to his supervisor about this thesis, and Erich ø to the dean about departmental policies.

BIBLIOGRAPHY

Bach, E. (1980), "Tenses and adverbs as functions on verb phrases," in C. Rohrer (ed.), Time, Tenses and Quantifiers, Max Niemeyer Verlag.

Baker, C.L. (1971), "Stress level and auxiliary behavior in English," Linguistic Inquiry 2, 167-181.

Baker, C.L. (1981), "Auxiliary - adverb word order," Linguistic Inquiry 12, 309-315.

Baltin, M. (1982), "A landing site theory of movement rules," Linguistic Inquiry 13, 1-38.

Bellert, I. (1977), "On semantic and distributional properties of sentential adverbs," Linguistic Inquiry 8, 337-350.

Bowers, J.S. (1969), "Adjectives and adverbs in English," Indiana University Linguistics Club.

Chomsky, N. (1970), "Remarks on nominalization," in R.A. Jacobs and P.S. Rosenbaum (eds.), Readings in English Transformational Grammar, Waltham, Mass., Ginn & Co., 184-221.

Dowty, D.R. (1978), "Governed transformations as lexical rules in a Montague grammar," Linguistic Inquiry 9, 393-426.

Dowty, D.R. (1979), Word Meaning and Montague Grammar, Dordrecht, D. Reidel.

Dowty, D.R. (1980), "Sentence adverbs and VP-adverbs revisited," paper presented at the annual meeting of the Linguistic Society of America.

Ernst, T. (1983), "More on adverbs and stressed auxiliaries," Linguistic Inquiry 14, 542-549.

Gazdar, G. (1981), "Unbounded dependencies and coordinate structure," Linguistic Inquiry 12, 155-186.

Gazdar, G. (1982), "Phrase structure grammar," in P. Jacobson and G. Pullum (eds.), The Nature of Syntactic Representation, D. Reidel, Dordrecht, 131-186.

Gazdar, G. and G. Pullum (1981), "Subcategorization, constituent order and the notion 'head'," in M. Moortgat, H.V.D. Hulst and T. Hoekstra (eds.), The Scope of Lexical Rules, Foris Publications, Dordrecht, 107-123.

Gazdar, G. and G. Pullum (1982), "GPSG: A theoretical synopsis," Indiana University Linguistics Club.

Gazdar, G., G. Pullum, and I. Sag (1982), "Auxiliaries and related phenomena in a restrictive theory of grammar," Language 58, 591-638.

Gazdar, G. and I. Sag (1981), "Passive and reflexives in phrase structure grammar," in J. Groenendijk, T. Janssen, and M. Stokhof (eds.), Formal Methods in the Study of Language, Mathematical Centre Tracts, Amsterdam, 131-152.

Greenbaum, S. (1969), Studies in English Adverbial Usage, Miami linguistic series, no. 5, Miami, University of Miami Press.

Jackendoff, R. (1972), Semantic Interpretation in Generative Grammar, Cambridge, Mass., MIT Press.

Jackendoff, R. (1977), $\bar{X}$ Syntax: A Study of Phrase Structure, Cambridge, Mass., MIT Press.

Jacobson, P. (1982), "Evidence for gaps," in P. Jacobson and G. Pullum (eds.), The Nature of Syntactic Representation, D. Reidel, Dordrecht, 187-228.

Jacobson, S. (1964), Adverbial Positions in English, Uppsala Dissertation, Stockholm: AB Studentbok.

Keyser, S.J. (1968), "Review of Sven Jacobson: Adverbial positions in English," Language 44, 357-374.

Klein, E. and I. Sag (1981), "Semantic type and control," paper presented at the annual meeting of the Linguistic Society of America.

Lakoff, G. (1965), On the Nature of Syntactic Irregularity, Mathematical Linguistics and Automatic Translation, Report NSF-16, Aiken Computation Laboratory of Harvard University.

Lakoff, G. (1970), "Adverbs and opacity: A reply to Stalnaker,"

Maling, J. (1976), "Notes on quantifier-postposing," Linguistic Inquiry 7, 708-718.

Postal, P. (1974), On Raising, Cambridge, Mass., MIT Press.

Prince, E. and G. Prince (1980), "On differentiating between Topicalization, Focus-Movement, and Yiddish-Movement," presented at the annual meeting of the Linguistic Society of America.

Pullum, G. (1982), "Syncategorematicity and English infinitival to," Glossa 16, 181-215.

Radford, A. (1981), Transformational Syntax, Cambridge, Cambridge University Press.

Sag, I. (1977), Deletion and Logical Form, Indiana University Linguistics Club.

Sag, I. (1978), "Floated quantifiers, adverbs, and extraction sites," Linguistic Inquiry 9, 145-150.

Sag, I. (1980), "A further note on floated quantifiers, adverbs, and extraction sites," Linguistic Inquiry 11, 255-257.

Sag, I. (1982), "Coordination, extraction, and Generalized Phrase Structure Grammar," Linguistic Inquiry 13, 329-335.

Schreiber, P. A. (1971), "Some constraints on the formation of English sentence adverbs," Linguistic Inquiry 2, 83-101.

Schreiber, P. A. (1972), "Style disjuncts and the performative analysis," Linguistic Inquiry 3, 321-47.